Applied Mathematics

Conference Board of the Mathematical Sciences

CBMS

Regional Conference Series in Mathematics

Number 134

Applications of Polynomial Systems

David A. Cox

with contributions by
Carlos D'Andrea
Alicia Dickenstein
Jonathan Hauenstein
Hal Schenck
Jessica Sidman

AMS American Mathematical Society

with support from the National Science Foundation

Applications of Polynomial Systems

Conference Board of the Mathematical Sciences

CBMS

Regional Conference Series in Mathematics

Number 134

Applications of Polynomial Systems

David A. Cox

with contributions by
Carlos D'Andrea
Alicia Dickenstein
Jonathan Hauenstein
Hal Schenck
Jessica Sidman

Published for the
Conference Board of the Mathematical Sciences
by the

AMS American Mathematical Society
Providence, Rhode Island

with support from the

National Science Foundation

NSF-CBMS Regional Conference in the Mathematical Sciences on Applications of Polynomial Systems held at Texas Christian University, June 4–8, 2018.

Partially supported by the National Science Foundation.
The author acknowledges support from the Conference Board of Mathematical Sciences and NSF grant number DSM 1741730. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

2010 *Mathematics Subject Classification.* Primary 13P15, 13P25, 14Q99; Secondary 14M25, 52C25, 62H99, 65H20, 68U07, 92C42.

For additional information and updates on this book, visit
www.ams.org/bookpages/cbms-134

Library of Congress Cataloging-in-Publication Data

Names: Cox, David A., author. | D'Andrea, Carlos (Carlos Antonio), 1973– author. | Dickenstein, Alicia, author. | Hauenstein, Jonathan D., author. | Schenck, Henry K., 1963– author. | Sidman, Jessica, author. | Conference Board of the Mathematical Sciences, issuing body. | National Science Foundation (U.S.), sponsoring body. | NSF-CBMS Regional Conference in the Mathematical Sciences (2018: Fort Worth, Tex.)

Title: Applications of polynomial systems / David A. Cox; with contributions by Carlos D'Andrea, Alicia Dickenstein, Jonathan Hauenstein, Hal Schenck and Jessica Sidman.

Description: Providence, Rhode Island: Published for the Conference Board of the Mathematical Sciences by the American Mathematical Society, [2020] | Series: Conference Board of the Mathematical Sciences CBMS regional conference series in mathematics, 0160-7642; Number 134 | "With support from the National Science Foundation." | Includes bibliographical references and index.

Identifiers: LCCN 2019045985 | ISBN 9781470451370 (paperback) | ISBN 9781470455897 (ebook)

Subjects: LCSH: Polynomials–Congresses. | Commutative algebra–Congresses. | Geometry, Algebraic–Congresses. | AMS: Commutative algebra – Computational aspects and applications – Solving polynomial systems; resultants. | Algebraic geometry – Computational aspects in algebraic geometry – None of the above, but in this section. | Commutative algebra – Computational aspects and applications – Applications of commutative algebra (e.g., to statistics, control theory, optimization, etc.). | Algebraic geometry – Special varieties – Toric varieties, Newton polyhedra . | Convex and discrete geometry – Discrete geometry – Rigidity and flexibility of structures. | Statistics – Multivariate analysis – None of the above, but in this section. | Numerical analysis – Nonlinear algebraic or transcendental equations – Global methods, including homotopy approaches. | Computer science – Computing methodologies and applications – Computer-aided de | Biology and other natural sciences – Physiological, cellular and medical topics – Systems biology, networks.

Classification: LCC QA161.P59 C69 2020 | DDC 512.9/422–dc23

LC record available at https://lccn.loc.gov/2019045985

---

10 9 8 7 6 5 4 3 2 1 25 24 23 22 21 20

Contents

Preface

Polynomials have been around for a long time. They play a prominent role in many applications and are basic objects in commutative algebra and algebraic geometry. In the 20th century, the foundations of algebraic geometry required a level of abstraction that made these tools hard to use for many people interested in applications. But with the advent of powerful algorithms, personal computers, and the theory of sparse polynomials, we now have *computational commutative algebra* and *computational algebraic geometry*. Coupled with accessible introductory texts, recent years have witnessed a remarkable range of applications.

The goal of this volume is to introduce some of these applications. The five chapters are based on the NSF-CBMS Regional Research Conference Applications of Polynomial Systems held at Texas Christian University (TCU), June 4–8, 2018. The format of the conference, replicated in this book, is that each day David Cox gave two lectures on a topic, followed by a third lecture by an expert on the topic.

The book is not a complete introduction to the topics covered. Many proofs are sketched or omitted, and the emphasis is on the examples. The hope is that the brief presentation provided here inspires you to learn more.

The first two chapters set the background for the rest of the book. Chapter 1 focuses on the rich history of elimination theory from the work of Newton and Bézout in the 17th and 18th centuries through the decline of elimination theory in the middle of the 20th century, followed by the emergence of symbolic computation in the latter part of the century. In the final section of the chapter, Carlos D'Andrea brings elimination theory into the 21st century.

Chapter 2 begins with the tension between the perfect information needed for symbolic computation and the messiness of the information coming from the real world. The chapter explores linear algebra and homotopy continuation, two commonly used methods for solving systems of polynomial equations numerically. This is where we meet our first substantial applications. Jonathan Hauenstein concludes the chapter with a discussion of sampling in numerical algebraic geometry.

The remaining chapters of the book focus on three substantial applications. Chapter 3 explores the geometry and algebra of geometric modeling, with some unexpected connections to toric varieties, algebraic statistics, and Rees algebras. At the end of the chapter, Hal Schenck surveys a variety of algebraic tools which are used in geometric modeling.

Chapter 4 is devoted to the geometry and combinatorics of rigidity theory. For bar-and-joint frameworks, basic objects include graphs and rigidity matrices, though polytopes and matroids also have an important role to play. A fun result due to Gross and Sullivant uses the rigidity matroid to study maximum likelihood estimates of Gaussian graphical models coming from planar graphs. Jessica Sidman

ends the chapter with a discussion of body-and-bar and body-and-cad frameworks and the study of non-generic frameworks via the Grassmann-Cayley algebra.

Chapter 5 shifts the scene to chemical reaction networks, where the Law of Mass Action leads to a lovely combination of graph theory and dynamical systems. This framework includes examples from chemistry, biology, epidemiology, and population genetics. Algebraic geometry enters the picture because the steady states are defined by a polynomial system. Toric varieties also have a role to play by the pioneering work of Karin Gatermann. Alicia Dickenstein finishes the chapter with an exposition of the interesting things that can happen in biochemical reaction networks.

Each chapter starts at a relatively elementary level, with more advanced topics introduced as needed. The algebraic geometry background can be found in the book [**104**]. In addition to the main applications presented in Chapters 3, 4, and 5, the twin themes of toric varieties and algebraic statistics play a prominent role in the book. Readers unfamiliar with this material should consult the expository papers [**98**] and [**214**].

We hope you enjoy the book! We had fun writing it.

Acknowledgments

The authors first want to thank the Conference Board of the Mathematical Sciences and the National Science Foundation for creating and supporting the NSF-CBMS Regional Research Conference Series. This book is a direct result of the NSF-CBMS conference held in June 2018, funded by NSF grant DMS-1741730 with additional support provided by the CBMS, the TCU Department of Mathematics, and the TCU College of Science and Engineering. Thank you all!

We are also tremendously grateful to the local organizers Greg Friedman, José Carrión, Eric Hanson, Scott Nollett, and Efton Park for their work in making the conference possible. They helped create a wonderful atmosphere for an intense week of mathematics. Thank you!

Final thanks go to the participants. The conference and the book benefited tremendously from your enthusiasm and excellent questions.

David Cox
I first want to thank the local organizers for inviting me to be principal lecturer for the CBMS conference they hoped to host. I have equal thanks for my wonderful coauthors, Carlos D'Andrea, Jonathan Hauenstein, Hal Schenck, Jessica Sidman, and Alicia Dickenstein, whose contributions enriched the conference and the book immensely.

I am also very grateful to Doug Arnold, Dan Bates, Alicia Dickenstein, Ron Goldman, Jonathan Hauenstein, Martin Kreuzer, Kaie Kubjas, Kuei-Nuan Lin, Nathan Pflueger, Henk Poulisse, Jan Peter Schäfermeyer, Gabriel Sosa, Tom Sederberg, Jessica Sidman, Frank Sottile, Simon Telen, Caroline Uhler, and Charles Wampler for their suggestions, comments, and assistance.

Finally, I want to acknowledge my indebtedness to Bernd Sturmfels for writing *Solving Systems of Polynomial Equations* [**334**]. This wonderful book, the result of a CBMS conference in 2002, is a model of vivid, concise writing that conveys the amazing ability of algebraic geometry to shed light on systems of polynomial equations.

Carlos D'Andrea
I am very grateful to Laurent Busé, Alicia Dickenstein, Emiliano Gomez, Martín Sombra, Teresa Krick, and Josephine Yu for their careful reading of an earlier version of this manuscript and several suggestions for improving the presentation.

Jonathan Hauenstein
I am very appreciative of discussions with Danielle Brake, Maggie Regan, and Sam Sherman to help formulate the presentation. I am also grateful for many conversations regarding sampling with Dan Bates, Heather Harrington, Chris Peterson, Andrew Sommese, Frank Sottile, Bernd Sturmfels, and Charles Wampler.

Hal Schenck
I thank David Cox—while I was a postdoc, David asked me a question on vanishing of syzygies, which led to a joint paper followed by a 20-year collaboration and lasting friendship. I'm also grateful to my coauthors Alicia Dickenstein, Alexandra Seceleanu, Javid Validashti, Eliana Duarte, and Nicolas Botbol.

Jessica Sidman
I am deeply indebted to Audrey St. John for introducing me to rigidity theory and am grateful to collaborators Zvi Rosen, Louis Theran, and Cynthia Vinzant, from whom I have learned so much. I would also like to thank all of the undergraduate contributors to projects mentioned here, including James Farre, Helena Kleinschmidt, Stephanie Stark, and Xilin Yu. Thanks also go to Bernd Sturmfels for suggesting computations in distance coordinates.

Alicia Dickenstein
I am very grateful to David Cox for including me in this project. I am also indebted to Bernd Sturmfels for introducing me to the subject of Chapter 5 while I was walking along a corridor at MSRI during the Commutative Algebra Program in the spring of 2003. I deeply thank Mercedes Pérez Millán, Nicolás Botbol, Magalí Giaroli, Frédéric Bihan, Anne Shiu, Elisenda Feliu, and David Cox for all I learned from them while working together.

David A. Cox, Amherst
Carlos D'Andrea, Barcelona
Jonathan Hauenstein, South Bend
Hal Schenck, Auburn
Jessica Sidman, South Hadley
Alicia Dickenstein, Buenos Aires

CHAPTER 1

Elimination Theory

Elimination theory has important roles to play in both algebraic geometry and symbolic computation. The first two sections of the chapter take a historical approach so that you can see how elimination theory developed in the last 300 years. In the third section Carlos D'Andrea will survey recent developments.

1.1. Elimination Theory in the 18th and 19th Centuries

Algebra was well-established in Europe by the end of the 17th century. The full story of how this happened and what followed in the subsequent two centuries is far beyond the scope of this volume. Readers interested in the details should consult [**17**] (for algebra in general) and [**287**] (for elimination theory). We will instead focus on some vignettes that illustrate the evolution of elimination theory.

Newton and Tschirnhaus. In spite of the title of the section, we begin in the 17th century with extracts from Newton and Tschirnhaus. In a manuscript written sometime between 1666 and 1671, Newton writes:

> *Datis duabus curvis invenire puncta intersectionis.* this is rather a principle then a probleme. But rather propounded of y^e Algebraicall then geometricall solutions & y^t is done by eliminating one of the two unknown quantitys out of y^e equations. From whence it will appeare y^t there are soe many cut points as the rectangle of the curves dimensions. [**282**, p. 177]

The Latin says "Given two curves, to find their points of intersection." The English is archaic, but once you realize that "y^e = the" and "y^t = that", Newton's meaning is clear:

- He solves the geometric problem by working algebraically.
- The algebra is "done by eliminating one of the two unknown quantitys out of y^e equations."
- The degree of the resulting equation is the product of the degrees of the curves ("the rectangle of the curves dimensions").

This is a clear statement of Bézout's Theorem, along with the strategy of reducing to a single variable that was to dominate the early history of elimination theory.

Here is an application of Newton's "principle."

Example 1.1. Newton counts the number of tangent lines to a curve that go through a given point:

> a line drawn from a given point may touch a curve of 2 dimensions in 1×2 points, of 3 in 2×3 points, of 4 in 3×4 points, &c. [**282**, p. 179]

We explain this as follows. Let $U(x,y,z)=0$ define a smooth curve of degree of n in the projective plane $\mathbb{P}^2$. If the given point is $P=(\alpha,\beta,\gamma)$, then any solution of

$$\begin{aligned} U &= 0\\ \alpha\frac{\partial U}{\partial x}+\beta\frac{\partial U}{\partial y}+\gamma\frac{\partial U}{\partial z} &= 0 \end{aligned}$$

is a point on the curve whose tangent line contains P. These curves have degrees n and $n-1$, giving $n(n-1)$ tangent lines by Bézout's Theorem. This explains the products 1×2, 2×3 and 3×4 in the quote. ◁▷

Another example of 17th century elimination can be found in the work of Tschirnhaus. In 1683, he published a paper on roots of a polynomial equation [**348**]. Given $y^3-qy-r=0$, Tschirnhaus introduces a new variable z via the *Tschirnhaus transformation* $y^2=by+z+a$, where a and b are unknown coefficients. He writes:

> following from this let there be a third equation (by proceeding according to the recognized rules of analysis) in which the quantity y is absent, and we shall obtain
>
> $$z^3+(3a-2q)z^2+(3a^2-4qa+q^2-qb^2+3rb)z+\cdots=0.$$

This leads to a solution of $y^3-qy-r=0$ as follows:

- Pick a to make the coefficient of z^2 vanish.
- Use the quadratic formula to pick b to make the coefficient of z vanish.
- The equation reduces to $z^3=$ constant, so z is a cube root.
- Solving $y^2=by+z+a$ for y gives a root of our original cubic.

Tschirnhaus's approach is based on elimination ("in which the quantity y is absent") using the "recognized rules of analysis."

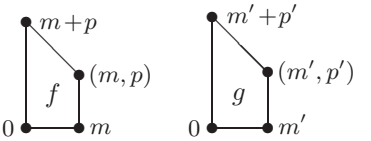
These examples show that elimination theory was an important part of algebra in the 17th century.

Cramer and Bézout. By the middle of the 18th century, Bézout's Theorem was well-known. Here is Cramer writing in 1750:

> If one has two variables and two indeterminate equations ... of which one is of order m and the other of order n, when, by means of these equations, one expels one of these variables, the one that remains has, in the final equation ... at most mn dimensions. [**116**]

In his 1764 paper *Sur le degré des équations résultantes de l'évanouissement des inconnues* [**27**], Bézout uses the terms "final equation" and "resultant equation" for the result of the elimination. The first main result of [**27**] is a version of Bézout's Theorem that draws on earlier work of Euler.

EXAMPLE 1.2. Let $f(x,y)$ and $g(x,y)$ have Newton polytopes:



Thus f has total degree $m+p$ in x,y, but its degree in x is bounded by m, with a similar story for g.

To eliminate x from $f(x,y)=g(x,y)=0$, Bézout multiplies f,g by polynomials G,H of degree $m'-1, m-1$ in x with unknown coefficients. Then equating the

coefficients of powers of x in the equation

$$Gf + Hg = h(y) \tag{1.1}$$

gives a linear system of $m + m'$ unknowns in $m + m'$ variables with a familiar coefficent matrix. We illustrate this for $m = m' = 2$, where

$$f(x, y) = A(y)x^2 + B(y)x + C(y),\ g(x, y) = A'(y)x^2 + B'(y)x + C'(y),$$

and $\deg(A(y)) \leq p$, $\deg(B(y)) \leq p + 1$, $\deg(C(y)) \leq p + 2$, with similar bounds for the degrees of $A'(y), B'(y), C'(y)$. If we set

$$G = Mx + N,\ H = M'x + N'$$

for unknown coefficients M, N, M', N', then equating coefficients of x in (1.1) gives the system of equations

$$\begin{bmatrix} A & 0 & A' & 0 \\ B & A & B' & A' \\ C & B & C' & B' \\ 0 & C & 0 & C' \end{bmatrix} \begin{bmatrix} M \\ N \\ M' \\ N' \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ h \end{bmatrix} \tag{1.2}$$

whose coefficient matrix is the Sylvester matrix, over 75 years before Sylvester wrote it down in 1840. Bézout goes on to show that when x is eliminated, the "équation résultant" in y has degree $mm' + mp' + m'p$. ◁▷

These days, we use the term *sparse elimination* for Example 1.2, yet Bézout's paper dates from 1764. We will soon see that Bézout had even more exotic versions of his theorem.

We turn to a second example from Bézout's paper.

Example 1.3. Suppose we have cubic equations

$$f = Ax^3 + Bx^2 + Cx + D = 0$$
$$g = A'x^3 + B'x^2 + C'x + D' = 0.$$

Then Bézout proceeds as follows:

(1) Multiply the first equation by A', the second by A, and subtract to obtain
$$(A'B - AB')x^2 + (A'C - AC')x + A'D - AD' = 0.$$
(2) Multiply the first equation by $A'x + B'$, the second by $Ax + B$, and subtract to obtain
$$(A'C - AC')x^2 + (A'D - AD' + B'C - BC')x + B'D - BD' = 0.$$
(3) Multiply the first equation by $A'x^2 + B'x + C'$, the second by $Ax^2 + Bx + C$, and subtract to obtain
$$(A'D - AD')x^2 + (B'D - BD')x + C'D - CD' = 0.$$

Bézout then considers "each power of x as an unknown," so that (1)–(3) give a system of three equations in three unknowns. The determinant of this system is the final equation that eliminates x. ◁▷

Example 1.3 and its generalization to $\deg(f) = \deg(g) = m$ (also studied by Bézout) is the origin of what we now call the *Bézoutian* of f, g. Observe that

$$y^2 \cdot \text{RHS of (1)} + y \cdot \text{RHS of (2)} + \text{RHS of (3)} = -\frac{f(x)g(y) - f(y)g(x)}{x - y},$$

which means that up to a minus sign, the quantities $A'B - AB', A'C - AC'$, etc. of Bézout are the coefficients b_{ij} in the expression

$$\frac{f(x)g(y) - f(y)g(x)}{x - y} = \sum_{i,j=1}^{3} b_{ij} x^{i-1} y^{j-1} \tag{1.3}$$

The *Bézoutian matrix* is $\mathrm{Bez}(f, g) = (b_{ij})$, and its determinant is the resultant $\mathrm{Res}(f, g)$. We will see later in the section that Cayley discovered this in 1857.

Bézout and Waring. We now turn to two important books:

- In 1779, Bézout published *Théorie Générale des Équations Algébriques*.
- In 1782, Waring published the second edition of *Meditationes Algebraicæ*.

We will go back and forth between the two books, beginning with a quote from Bézout's *Théorie Générale* (italics added):

> We conceive of each *given equation* as being multiplied by a *special polynomial*. Adding up all those products together, the result is what we call the *sum-equation*. This sum-equation will become the *final equation* through the vanishing of all terms affected by the unknowns to eliminate. [**28**, (224.)]

Here are the terms in italics, expressed in modern language:

- We have *given equations* $f_1 = \cdots = f_n = 0$ in $x_1, \ldots, x_n$ (for Bézout, the number of equations equals the number of unknowns).
- Multiply each by *special polynomials* $A_1, \ldots, A_n$.
- Adding then up gives the *sum-equation* $A_1 f_1 + \cdots + A_n f_n = 0$.
- This becomes the *final equation* when all terms involving the unknowns to eliminate vanish, e.g.,

$$A_1 f_1 + \cdots + A_n f_n \in k[x_1].$$

Thus the elimination ideal $\langle f_1, \ldots, f_n \rangle \cap k[x_1]$ is implicit in Bézout. But set theory lay over 100 years in the future, so objects like $\langle f_1, \ldots, f_n \rangle$ were not part of Bézout's mental landscape—he worked with polynomials finitely many at a time. Even writing $k[x_1]$ is problematic, for Bézout never specified a field. His definition of polynomial simply says "coéfficiens quelconques" ("any coefficients").

Now let's switch to Waring's *Meditationes Algebraicæ*. This book covers an impressive range of topics, from symmetric polynomials to the Waring problem about expressing positive integers as sums of kth powers. We will focus on Waring's comments about Bézout's *Théorie Générale*, which appeared three years before the second edition of *Meditationes Algebraicæ* [**358**].

Here is a quote from Waring, where we have added bullets and comments enclosed in [...] that indicate the relation to Bézout:

> If there are h equations [*given equations*], of degrees n, m, l, k, etc., respectively, in as many unknowns, and
>
> - if each of these h equations are multiplied by h assumed equations [*special polynomials*] of degrees $nmlk\cdots - n$, $nmlk\cdots - m$, $nmlk\cdots - l$, $nmlk\cdots - k$, etc., respectively, then
> - let the resulting equations be added together [*sum-equation*], and
> - equate to zero the coefficients of each term of degree $nmlk\cdots$ in the result (barring those involving only powers of y),
>
> whence the equation whose root is x or y or z, etc. [*final equation*], cannot be of degree greater than $nmlk\cdots$ [**358**, p. xxxv, English translation]

This is Bézout's Theorem for n polynomial equations in n unknowns, cast as a problem in elimination theory. We will state Bézout's version below.

Here is another of Bézout's results from *Théorie Générale*, again quoting from Waring with bullets added:

> If there are h equations respectively of degrees $n, m, l, k, \ldots$,
> - all involving the same unknown quantities $x, y, z, v, \ldots$, and
> - if $p, q, r, s, \ldots, p', q', r', s', \ldots, p'', q'', r'', s'', \ldots$, etc., are the maximum degrees to which $x, y, z, v, \ldots$ appear in the equations of degrees $n, m, l, k, \ldots$ respectively,
> - then the equation whose root is x or y or z, etc., cannot be of degree higher than
>
> $$\begin{aligned} n \times m \times l \times k \times \cdots &- (n-p) \times (m-p') \times (l-p'') \times \cdots \\ &- (n-q) \times (m-q') \times (l-q'') \times \cdots \\ &- (n-r) \times (m-r') \times (l-r'') \times \cdots \\ &- \cdots \end{aligned}$$
>
> [**358**, p. 209, English translation]

The key feature here is that we fix the total degree of each polynomial and put separate bounds on the degree in each variable. This differs from Example 1.2, where in 1764 Bézout fixed the total degree and bounded the degree of just the first variable.

Bézout's *Théorie générale.* The results quoted from Waring are just the tip of the iceberg when it comes to the "Bézout theorems" in *Théorie Générale* [**28**]. This amazing book consists of 469 pages of text organized into sections (1.)–(561.).

Bézout had an elaborate classification of polynomials and systems of equations. He worked affinely with polynomials of degree T in n variables. A polynomial is *complete* when it uses *all* monomials of total degree $\leq T$. We are now ready for Bézout's first theorem:

> *The degree of the "équation finale résultant" of any number of complete equations containing the same number of unknowns, and of any degree, is equal to the product of the exponents of the degrees of these equations.* [**28**, (47.)].

To relate this to the number of solutions, recall that the "équation finale résultant" comes from eliminating all variables but one, say x_1, so its roots are the x_1-coordinates of solutions of the system. Hence, if the number of solutions is finite with distinct x_1 coordinates, then the number of solutions is bounded by the "product of the exponents of the degrees" since some solutions might be at ∞ (recall that we are working affinely).

For Bézout, a polynomial is *incomplete* when it is not complete. Incomplete polynomials fall into various *species*. For example, the second result of Bézout quoted from Waring is for polynomials of the *first species*. To indicate how deeply Bézout thought about polynomials, let's explore his third species in detail.

EXAMPLE 1.4. A polynomial in u, x, y belongs to the third species considered by Bézout when it is satisfies the following conditions:

1.° that u does not exceed the degree a, x does not exceed the degree $\underset{,}{a}$, y does not exceed the degree $\underset{,,}{a}$;

2.° that u with x does not rise above the dimension b; u with y does not rise above the dimension $\underset{,}{b}$; x with y does not rise above the dimension $\underset{,,}{b}$;

3.° that u with x and with y cannot together rise to a dimension higher than t. [**28**, (82.)]

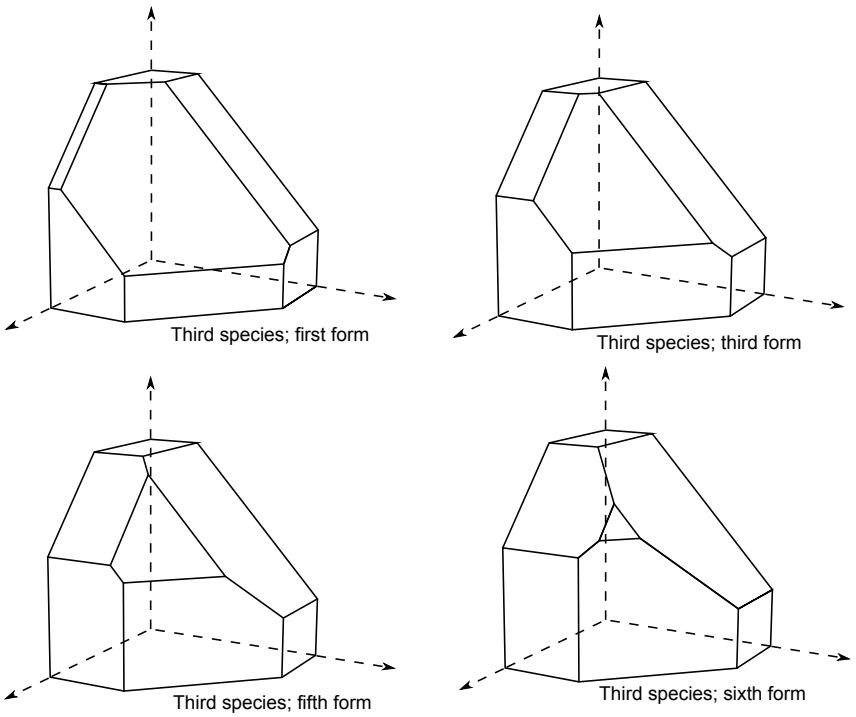
Then a monomial $u^{k_1}x^{k_2}y^{k_3}$ is of the third species considered by Bézout if there are constants $t, a, \underset{'}{a}, \underset{''}{a}, b, \underset{'}{b}, \underset{''}{b}$ (in Bézout's awkard notation) such that

$$\begin{aligned} &k_1 + k_2 + k_3 \le t \\ &0 \le k_1 \le a, \quad 0 \le k_2 \le \underset{'}{a}, \quad 0 \le k_3 \le \underset{''}{a} \\ &k_1 + k_2 \le \underset{''}{b}, \quad k_1 + k_3 \le \underset{'}{b}, \quad k_2 + k_3 \le b. \end{aligned} \tag{1.4}$$

Bézout notes that the third species has eight *forms*, according to the signs of

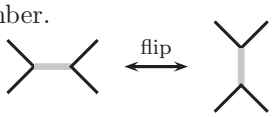
$$t - b - \underset{'}{b} + \underset{''}{a}, \quad t - b - \underset{''}{b} + \underset{'}{a}, \quad t - \underset{'}{b} - \underset{''}{b} + a. \tag{1.5}$$

These days, we regard the inequalities (1.4) as describing the lattice points in a polytope, and the eight forms indicate give eight possibilities for what the polytope can look like. Here is a drawing of four of these, taken from [**288**]:



To a reader versed in toric geometry, this picture is astonishing. The polytopes share the same 10 inward-facing facet normals. These 10 vectors have a *secondary fan* whose chambers have some wonderful properties:

- Each polytope in the picture corresponds to a chamber.
- Two chambers that share a common wall give polytopes that differ by a *bistellar flip* of an edge:



In the picture, you can see a sequence of flips about edges:

$$\text{first form} \overset{\text{flip}}{\longleftrightarrow} \text{third form} \overset{\text{flip}}{\longleftrightarrow} \text{fifth form} \overset{\text{flip}}{\longleftrightarrow} \text{sixth form}.$$

This corresponds to a path in secondary fan where we traverse four chambers, crossing three walls in the process. See [**106**, §14.4] for more on the secondary fan.

However, Bézout was not thinking about the splendid geometry implicit in the eight forms of the third species. Rather, he wanted the degree of the "équation finale résultant" of a system $f = g = h = 0$ when the polynomials have the same form of the third species. In [**28**, (120.)–(127.)], he states eight Bézout Theorems when f, g, h satisfy (1.4) for parameters

$$(t, a, \underset{,}{a}, \underset{,,}{a}, b, \underset{,}{b}, \underset{,,}{b}),\ (t', a', \underset{,}{a}', \underset{,,}{a}', b', \underset{,}{b}', \underset{,,}{b}'),\ (t'', a'', \underset{,}{a}'', \underset{,,}{a}'', b'', \underset{,}{b}'', \underset{,,}{b}''). \tag{1.6}$$

If in addition f, g, h are all of the first form, where the quantities in (1.5) are all negative, then the "équation finale résultant" has degree

$$\begin{aligned} D = {} & tt't'' - (t-a).(t'-a').(t''-a'') - (t-\underset{,}{a}).(t'-\underset{,}{a}').(t''-\underset{,}{a}'') \\ & - (t-\underset{,,}{a}).(t'-\underset{,,}{a}').(t''-\underset{,,}{a}'') + (t-b).(t'-b').(t''-b'') \\ & + (t-\underset{,}{b}).(t'-\underset{,}{b}').(t''-\underset{,}{b}'') + (t-\underset{,,}{b}).(t'-\underset{,,}{b}').(t''-\underset{,,}{b}'') \\ & - (a+\underset{,}{a}-b).(t'-b').(t''-b'') - (a'+\underset{,}{a}'-b').(t-b).(t''-b'') \\ & - (a''+\underset{,}{a}''-b'').(t-b).(t'-b') - (a+\underset{,,}{a}-\underset{,}{b}).(t'-\underset{,}{b}').(t''-\underset{,}{b}'') \\ & - (a'+\underset{,,}{a}'-\underset{,}{b}').(t-\underset{,}{b}).(t''-\underset{,}{b}'') - (a''+\underset{,,}{a}''-\underset{,}{b}'').(t-\underset{,}{b}).(t'-\underset{,}{b}') \\ & - (\underset{,}{a}+\underset{,,}{a}-\underset{,,}{b}).(t'-\underset{,,}{b}').(t''-\underset{,,}{b}'') - (\underset{,}{a}'+\underset{,,}{a}'-\underset{,,}{b}').(t-\underset{,,}{b}).(t''-\underset{,,}{b}'') \\ & - (\underset{,}{a}''+\underset{,,}{a}''-\underset{,,}{b}'').(t-\underset{,,}{b}).(t''-\underset{,}{b}'), \end{aligned} \tag{1.7}$$

where we follow the typsetting used in [**28**, (120.)]. Bézout gives similar formulas when the polynomials all have the second form, the third form, etc.

Here's what's happening from the modern point of view. Assume that the parameters (1.6) of $f, g, h \in \mathbb{C}[u, x, y]$ are of the first form and that f, g, h are generic within their Newton polytopes. Then:

- The above formula for D is the *mixed volume* of the Newton polytopes of the polynomials f, g, h. We will verify this in Section 1.2.
- By Bernstein's Theorem (see Theorem 1.12 in Section 1.2), D is the number of solutions of $f = g = h = 0$ in the torus $(\mathbb{C}^*)^3$.
- The minimal generator of the elimination ideal $\langle f, g, h\rangle \cap \mathbb{C}[u]$ has degree D. This polynomial is Bézout's "équation finale résultant."

It is wonderful to see so much contemporary mathematics packed into an example from a book published in 1779. ◁▷

Bézout's Proofs. In [**288**], Penchèvre gives a careful reading of *Théorie générale* and discusses Bézout's proofs, most of which are incomplete. But [**288**] also shows how tools from commutative algebra (Koszul complexes) and algebraic geometry (toric varieties) can be used to verify many Bézout's results, including the formula for D given in Example 1.4.

To give you a hint of what Bézout did, let's say a bit more about Example 1.4. If we set $\mathsf{a} = (a, \underset{,}{a}, \underset{,,}{a})$, then the data (1.6) can be written

$$(t, \mathsf{a}, \mathsf{b}),\ (t', \mathsf{a}', \mathsf{b}'),\ (t'', \mathsf{a}'', \mathsf{b}'')$$

To find the "équation finale résultant," Bézout uses *special polynomials* F, G, H so that $Ff + Gg + Hg$ lies in the elimination ideal. He picks F to have parameters

$(T-t, \mathsf{A}-\mathsf{a}, \mathsf{B}-\mathsf{b})$ so that Ff has parameters $(T, \mathsf{A}, \mathsf{B})$, and similarly for G, H. In modern language, this leads to the linear map

$$\phi : C_1 \oplus C_2 \oplus C_3 \xrightarrow{(f,g,h)} C_0,$$

where all possible F's lie in C_1, etc. Bézout is interested in $\dim \operatorname{coker} \phi$, which he describes by saying *useful coefficients* and *useless coefficients*. Then:

- The dimension of C_0 is the number of lattice points in the polytope (1.4) for parameters $(T, \mathsf{A}, \mathsf{B})$, and similarly for C_1, C_2, C_3.
- This number is a quasi-polynomial in the parameters, with a separate polynomial for each of the eight forms. Bézout computes these in [**28**, (92.)–(99.)].
- Bézout's formula for $\operatorname{coker} \phi$ implicitly assumes that the Koszul complex of f, g, h is exact, which means that the only cancellations that occur are the obvious ones. Bézout explains this by saying "the number of terms cannot be made lower by introducing fictitious terms." [**28**, (112.)]

All of this is discussed in detail in [**288**, Section 6]. There is a lot of math going in Bézout's book!

Poisson. While a student at the École Polytechnique, Poisson wrote the paper *Mémoire sur l'élimination dans les équations algébriques*, published in 1802. Here is how his paper begins:

> The degree of the "équation finale résultant" from the elimination, among a number of equations equal to that of the unknowns, of all the unknowns except one, cannot be greater than the product of the exponents of these equations; and it is precisely equal to this product, when the given equations are the most general of their degrees. This important theorem is Bézout's, but the way he proves it is neither direct nor simple; nor is it devoid of any difficulty. [**293**]

Poisson's goal is to prove Bézout's Theorem in the case of complete polynomials, and he observes that Bézout's bound is sharp in the generic case ("most general of their degrees"). The final sentence of the quote is telling—like many of his contemporaries, Poisson found Bézout's book to be difficult reading. This happened so often that in 1907, Brill and Noether [**45**, p. 143] would comment that Bézout's *Théorie générale* is "ebenso viel berühmten als wenig gelesen" ("as well known as little read").

Because of this, Poisson provides his own proof, which proceeds by induction on n, the number of equations in the system $f_1 = \cdots = f_n = 0$. The inductive step uses the product

$$\prod_{f_1(p)=\cdots=f_{n-1}(p)=0} f_n(p),$$

which is now part of the *Poisson formula* for the multivariable resultant (see, for example, [**105**, Exercise 3.3.8]). For the base case $n = 2$ of the induction, Poisson cites Cramer's 1750 paper [**116**], where Cramer uses the product $\prod_{f_1(p)=0} f_2(p)$.

Poisson's paper brings us to the end of our discussion of elimination in the 18th century. Readers interested in a fuller picture of what happened in this century should consult Penchèvre's article [**286**].

The 19th Century. The century that followed Poisson's 1802 paper witnessed an explosion of work on elimination theory. In 1892, Netto and Le Vavasseur wrote a 232 page review article about algebra in the 19th century for *l'Encyclopédie des*

sciences mathématiques pures et appliquées. Their section on elimination theory is 97 pages long and begins as follows:

> The large number and variety of memoirs relating to elimination makes it difficult to classify these memoirs rationally. [**281**, pp. 73–169]

A more recent account of elimination theory in the 19th century can be found in Penchèvre's 318 page PhD thesis *Histoire de la théorie de l'élimination* [**287**, Chapters 12–16].

There is no way we can do justice to this vast amount of material. We will instead focus on a few selected topics, beginning with resultants.

The Emergence of Resultants. After Poisson, elimination theory had a low profile in the first few decades of the 19th century. In 1835, Jacobi [**218**] used elimination theory to prove a result that became the *Euler-Jacobi Formula* (see [**245**] for a modern account). But starting in 1839, multiple papers on elimination theory appeared in rapid succession.

The Sylvester matrix appeared twice during this period. In 1840, Sylvester [**340**] described a matrix whose rows are built from the coefficients of the polynomials, suitably shifted by successively adding more zeros. The result is

> a solid square $(m+n)$ terms *deep* and $(m+n)$ terms *broad*.

He takes the determinant of this matrix, though he never says "determinant."

The same matrix was discovered independently four years later by Hesse in 1844 [**204**]. He begins with polynomials

$$A_0 = a_n\,x^n + a_{n-1}\,x^{n-1} + a_{n-2}\,x^{n-2} + \ldots. + a_1x + a_0$$
$$B_0 = b_mx^m + b_{m-1}x^{m-1} + b_{m-2}x^{m-2} + \ldots. + b_1x + b_0$$

(he uses four dots .... rather than the three dots ... commonly used these days). He multiplies A_0 by $x^{m-1}, x^{m-2}, \ldots, x, 1$ and B_0 by $x^{n-1}, x^{n-2}, \ldots, x, 1$. Expressing these in terms of the monomials $x^{m+n-1}, x^{m+n-2}, \ldots, x, 1$, he obtains the following $(m+n)\times(m+n)$ matrix:

	1	2	3	....	....	....	....	....	....	....	$m{+}n{-}1$	$m{+}n$
1	a_n	a_{n-1}	a_{n-2}	....	a_2	a_1	a_0	0	0	0	0	0
2	0	a_n	a_{n-1}	....	a_3	a_2	a_1	a_0	0	0	0	0
3	0	0	a_n	....	a_4	a_3	a_2	a_1	a_0	0	0	0
$\vdots$												
$m{-}1$	0	0	0	0	a_n	a_{n-1}	a_{n-2}	....	a_2	a_1	a_0	0
m	0	0	0	0	0	a_n	a_{n-1}	....	a_3	a_2	a_1	a_0
$m{+}1$	b_m	b_{m-1}	b_{m-2}	....	b_2	b_1	b_0	0	0	0	0	0
$m{+}2$	0	b_m	b_{m-1}	....	b_3	b_2	b_1	b_0	0	0	0	0
$m{+}2$	0	0	b_m	....	b_4	b_3	b_2	b_1	b_0	0	0	0
$\vdots$												
$m{+}n{-}1$	0	0	0	0	b_m	b_{m-1}	b_{m-2}	....	b_2	b_1	b_0	0
$m{+}n$	0	0	0	0	0	b_m	b_{m-1}	....	b_3	b_2	b_1	b_0

Like Sylvester, he arranged the coefficients in rows, so this is the transpose of what is commonly called the "Sylvester matrix." We saw an example of this in (1.2), though neither Sylvester nor Hesse seem to be aware of what Bézout did in 1764.

These papers from 1840 and 1844 worked affinely and said "result of elimination" rather than "resultant." But in 1847, Cayley published a paper on elimination that begins as follows:

> Designating $U, V, W, \ldots$ homogeneous functions of orders $m, n, p,$ &c. and an equal number of variables respectively, and assuming that these functions are the most general possible, that is to say that the coefficient of each term is an indeterminate letter: we know that the equations $U = 0, V = 0, W = 0, \ldots$ offer a relation $\Theta = 0$, in which the variables no longer enter, and where the function Θ, which can be called the *complete Resultant of the equations*, is homogeneous and of the order $np \ldots$ relative to the coefficients of U, of the order $mp \ldots$ relative to those of V, and so on, while it is not decomposable into factors. [**71**]

There is a *lot* going in this quote:

- Cayley explicitly says "resultant."
- He uses homogeneous polynomials. We are no longer affine!
- He states some basic properties of the resultant (degree in the coefficients of each polynomial, irreducibility).

We have the beginnings of a general theory of resultants in 1847!

The switch to homogeneous polynomials was driven by two emerging areas of mathematics: projective algebraic geometry, where points in projective space are described by homogeneous coordinates, and invariant theory, where invariants are naturally homogeneous (an example is the determinant, invariant under the action of $\mathrm{SL}(n)$). To get a better sense of how the story unfolded, let's spend some more time with Cayley.

Cayley. In 1845 and 1846, Cayley wrote two papers on invariant theory that introduced *hyperdeterminants*. In the second paper, he observes a link with the resultant of two homogeneous polynomials of degree two and comments that

> Important results might be obtained by connecting the theory of hyperdeterminants with that of elimination, but I have not yet arrived at anything satisfactory upon this subject. [**70**]

We suspect that Cayley would have enjoyed the 1994 book *Discriminants, Resultants, and Multidimensional Determinants* [**165**], which devotes an entire chapter to hyperdeterminants.

Cayley was also familiar with Bézout's 1764 paper. In 1857, he published the short paper *Note sur la méthode d'élimination de Bezout* in Crelle's journal [**72**][1] about what he later called "Bezout's abbreviated process of elimination." We illustrated this in Example 1.3. Cayley and Sylvester wrote Bézout's name without the accent.

In his note, Cayley begins with homogenous polynomials

$$U = (a, \ldots \text{)}x, y)^n = 0$$
$$V = (a', \ldots \text{)}x, y)^n = 0,$$

[1]When Cayley published outside of England, he often wrote in French—in the 19th century, English was not the dominant mathematical language that it is today.

where $(a,b,c,\dots)(x,y)^n = ax^n + \binom{n}{1}bx^{n-1}y + \binom{n}{2}cx^{n-2}y^2 + \cdots$ is how Cayley writes a homogeneous polynomial of degree n in x, y. He refers to U and V as "quantics." This is part of the symbolic notation that Cayley developed for invariant theory.

Then Cayley considers the quotient

$$\frac{(a,\dots)(x,y)^n\,(a',\dots)(\lambda,\mu)^n - (a,\dots)(\lambda,\mu)^n\,(a',\dots)(x,y)^n}{\mu x - \lambda y},$$

which we would write more simply as

$$\frac{U(x,y)V(\lambda,\mu) - U(\lambda,\mu)V(x,y)}{\mu x - \lambda y}. \tag{1.8}$$

As noted after Example 1.3, Bézout's method is encoded in the coefficients of (1.8) and the determinant of the resulting square matrix is the resultant of U and V. Also observe that (1.8) is the homogeneous version of what we did in (1.3).

A year later in his *Fourth Memoir Upon Quantics*, Cayley calls (1.8) the *Bezoutic emanant* and comments that

> the result of the elimination is consequently obtained by equating to zero the determinant formed with the matrix which enters into the expression of the Bezoutic emanant. In other words, this determinant is the Resultant of the two quantics. [**73**]

These days, we use the simpler term "Bézoutian," for which we should be grateful.

Sylvester also studied Bézout's work and in 1853 wrote a paper [**341**] that defines *Bezoutians*, *Bezoutics* and the *Bezoutian matrix*. This paper also introduced *Sylvester double sums*, a subject of current interest—see [**118**].

We next discuss a geometric example from Cayley that involves elimination.

EXAMPLE 1.5. In Example 1.1, Newton counted tangent lines to a curve that go through a given point. In his 1847 paper, Cayley poses the problem of finding the *equations* of these tangent lines:

> Find the equation of the system of tangents drawn from a fixed point to a given curve. [**71**]

As before, we assume that $U(x,y,z) = 0$ defines a smooth curve of degree of n in $\mathbb{P}^2$ and that the given point is $P = (\alpha,\beta,\gamma)$. In order for a point $Q = (\xi,\eta,\zeta)$ to lie on a tangent line going through P, we must satisfy the system of equations

$$\begin{aligned} U &= 0 \\ \alpha\frac{\partial U}{\partial x} + \beta\frac{\partial U}{\partial y} + \gamma\frac{\partial U}{\partial z} &= 0 \\ \det\begin{bmatrix} \xi & \eta & \zeta \\ x & y & z \\ \alpha & \beta & \gamma \end{bmatrix} &= 0. \end{aligned} \tag{1.9}$$

The first two equations tell us that (x,y,z) is a point on the curve whose tangent line goes through P, and the third equation says that Q lies on this tangent.

For this system, Cayley regards the variables as x, y, z, while $\alpha, \beta, \gamma, \xi, \eta, \zeta$ are unknown coefficients. He then takes the resultant Θ of (1.9), where Θ is from the quote at the top page 10. You should reread this quote before proceeding further.

The resultant Θ is a polynomial in $\alpha, \beta, \gamma, \xi, \eta, \zeta$ and the coefficients of U. He notes that (1.9) has an obvious solution when (α,β,γ) is on the curve (courtesy of

the Euler relation $xU_x + yU_y + zU_z = nU$), so that $U(\alpha, \beta, \gamma)$ is a factor of Θ. Thus

$$\Theta = U(\alpha, \beta, \gamma) \cdot \Phi,$$

where Φ is the defining equation of the tangent lines when regarded as a polynomial in ξ, η, ζ. In Cayley's terminology, Φ is a *reduced resultant*. ◁▷

Although the resultant is irreducible when the coefficients of the system are independent unknowns, *extraneous factors* can arise for the resultant of a particular system, such as (1.9). We will see some interesting examples of extraneous factors in Chapter 3.

We close our discussion of Cayley with a quote from a paper on elimination he wrote in 1864:

> In the problem of elimination, one seeks the relationship that must exist between the coefficients of a function or system of functions in order that some particular circumstance (or singularity) can occur. [**74**]

It is clear that for Cayley, elimination is at the heart of algebraic geometry.

Before we move on to the general theory of resultants, we need to discuss an interesting 1841 paper of Minding that illustrates a different flavor of elimination.

Minding. While the 19th century focused mostly on resultants of complete polynomials (in Bézout's terminology), we have seen that Bézout himself studied more general situations. But Minding's 1841 paper *Ueber die Bestimmung des Grades einer durch Elimination hervorgehenden Gleichung* [**275**] goes far beyond anything Bézout could have imagined. Minding considers a system of (affine) equations

$$\begin{aligned} f(x,y) &= A_0 y^m + A_1 y^{m-1} + \cdots + A_{m-1} y + A_m = 0 \\ \theta(x,y) &= B_0 y^n + B_1 y^{n-1} + \cdots + B_{n-1} y + B_n = 0, \end{aligned}$$

where "the letters A and B with subscripts stand for arbitrary polynomials in x." The goal is to compute the degree of the "final equation" $\psi(x) = 0$ obtained by eliminating y. Following Poisson, Minding expresses the final equation as

$$\psi(x) = B_0^m f(x, y_1) \cdot f(x, y_2) \cdots f(x, y_n), \tag{1.10}$$

where $y_1, \ldots, y_n$ are the roots of $\theta(x, y)$. Minding expresses each y_i as a Puiseux series in x about ∞, so the exponents are decreasing. For example,

$$y_1 = y_1(x) = c_1 x^{h_1} + d_1 x^{h_1 - \alpha_1} + e_1 x^{h_1 - \alpha_2} + \cdots, \quad 0 < \alpha_1 < \alpha_2 < \cdots,$$

where $h_1, \alpha_1, \alpha_2, \cdots \in \mathbb{Q}$. Doing this for $y_2, \ldots, y_n$ and focusing on the highest power of x on each side of (1.10), we obtain

$$\deg(\psi(x)) = \deg\big(B_0(x)^m f(x, c_1 x^{h_1}) \cdots f(x, c_n x^{h_n})\big).$$

If we set $b = \deg\big(B_0(x)\big)$ (an integer) and $k_i = \deg\big(f(x, c_i x^{h_i})\big)$ (often a fraction), then the above equation simplifies to

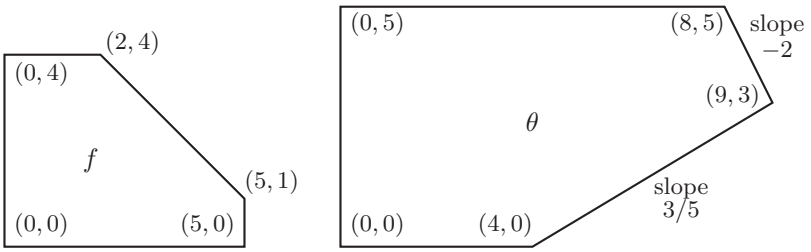
$$\deg(\psi(x)) = mb + k_1 + \cdots + k_n,$$

which is the major result of [**275**]. Here is an example from Minding's paper.

Example 1.6. Suppose that

$$\begin{aligned} f(x,y) &= (x^2)y^4 + (x^2)y^3 + (x^4)y^2 + (x^5)y + (x^5) \\ \theta(x,y) &= (x^8)y^5 + (x^6)y^4 + (x^9)y^3 + (x^4)y^2 + (x^3)y + (x^4), \end{aligned}$$

where (x^ℓ) means a polynomial of degree ℓ. These polynomials have the following Newton polygons:



The commentary to [**275**, English translation] explains how the negative reciprocals of the slopes of the polygon for θ give

$$h_1 = h_2 = \tfrac{1}{2},\ h_3 = h_4 = h_5 = -\tfrac{5}{3}.$$

This easily leads to

$$k_1 = k_2 = \tfrac{11}{2},\ k_3 = k_4 = k_5 = 5,$$

giving $\deg(\psi(x)) = 4 \cdot 8 + \frac{11}{2} + \frac{11}{2} + 5 + 5 + 5 = 58$. ◁▷

The commentary to [**275**, English translation] goes on to show that in general, Minding's formula for $\deg(\psi(x))$ equals the mixed volume (mixed area in this case) of the Newton polygons of f and θ. So Minding's paper from 1841 provides another splendid example of Bernstein's Theorem.

The Theory of Resultants. In the latter part of the 19th century, many papers on resultants were published, including Brill (1880) [**43**], Kronecker (1882) [**244**] and Mertens (1886) [**271**, **272**]. This continued into the early part of the 20th century with Macaulay (1902) [**264**] and the books by Netto (1900) [**280**] and Macaulay (1916) [**265**].

These papers and books focus primarily on homogeneous polynomials where all monomials of a given degree are allowed. This is the homogeneous analog of Bézout's notion of a complete polynomial. Let $f_0, \cdots, f_n$ be homogeneous in $x_0, \ldots, x_n$ of degrees $d_0, \ldots, d_n$. Then the *classical* or *dense resultant*, denoted by $\mathrm{Res}_{d_0,\ldots,d_n}(f_0, \ldots, f_n)$, is an integer polynomial in the coefficients of the f_i with the property that over any algebraically closed field,

$$\mathrm{Res}_{d_0,\ldots,d_n}(f_0, \ldots, f_n) = 0 \iff \begin{cases} f_0 = \cdots = f_n = 0 \text{ has a} \\ \text{solution in projective space } \mathbb{P}^n. \end{cases} \tag{1.11}$$

Kronecker, in his 1882 paper *Grundzüge einer arithmetischen Theorie der algebraischen Grössen* on the foundations of algebraic geometry, sketched a theory of resultants that gave elimination a central role. Here is a quote from his paper:

> A system of arbitrarily many algebraic equations for $\mathfrak{z}^0, \mathfrak{z}', \mathfrak{z}'', \mathfrak{z}^{(n-1)}$, in which the coefficients belong to the rationality domain $(\mathfrak{R}', \mathfrak{R}'', \ldots)$, define the algebraic relations between $\mathfrak{z}$ and $\mathfrak{R}$, whose knowledge and representation are the purpose of the theory of elimination. [**244**, §10]

Kronecker's notation and terminology are challenging for the modern reader—see [**289**] for a modern account of [**244**]. But what Kronecker is saying in the quote does not differ greatly from the 1864 quote from Cayley given earlier. Kronecker never published the details of his approach, leaving it to his student Mertens to complete the task.

The first book to cover resultants in detail was Volume II of Netto's *Vorlesungen über Algebra* [**280**], published in 1900. His treatment of resultants includes sections on Poisson, Minding, Kronecker, Cayley, Sylvester, and Bézout. He actually read Bézout's *Théorie Générale*!

In [**280**, pp. 146–147], Netto generalizes the matrix described by Sylvester and Hesse in the 1840s. Recall that when $f(x), g(x)$ have degree n, m, Hesse multiplied f by $x^{m-1}, \dots, x, 1$ and g by $x^{n-1}, \dots, x, 1$, and expressed these in terms of $x^{m+n-1}, \dots, x, 1$. This gave the Sylvester matrix whose determinant is $\operatorname{Res}(f, g)$.

Netto studies what happens in general. He works affinely, so we will modify his notation to work homogeneously. Consider $m+1$ homogeneous equations

$$f_\alpha(x_0, \dots, x_m) = 0, \quad \alpha = 0, \dots, m$$

of degree $n_\alpha = \deg(f_\alpha)$. We assume that f_α is the sum of all monomials of degree n_α with coefficients that are independent variables. This gives the universal version of the resultant in degrees $n_0, \dots, n_m$, which we denote by

$$\operatorname{Res}(f_0, \dots, f_m) = \operatorname{Res}_{n_0,\dots,n_m}(f_0, \dots, f_m) \in \mathbb{Z}[c].$$

Here, c is the vector of coefficients of $f_0, \dots, f_m$, so $f_0, \dots, f_m \in \mathbb{Z}[c, x_0, \dots, x_m]$.

To build a matrix that captures the resultant, set

$$n = n_0 + \cdots + n_m - m$$

and multiply f_α by all monomials x^β of degree $n - n_\alpha$. Then express the products $x^\beta f_\alpha$ as linear combinations of all monomials of degree n. The coefficients form the rows of a matrix M of size $\Sigma \times N$, where Σ is the number of products $x^\beta f_\alpha$ and N is the number of monomials z^β. Using the definition of n, Netto shows that

$$\Sigma = \sum_{\alpha=1}^{m+1} \binom{n - n_\alpha + m}{m} \geq N = \binom{n+m}{m}.$$

When $m = 1$, this is an equality and M is the Sylvester matrix of f_0 and f_1.

PROPOSITION 1.7. *Let Δ be an $N \times N$ maximal minor of M. Then:*

(1) Δ *is a multiple of* $\operatorname{Res}(f_0, \dots, f_m)$.
(2) $\Delta \in \mathbb{Z}[c] \cap \big(\langle f_0, \dots, f_m\rangle : \langle x_0, \dots, x_m\rangle^\infty\big)$.

PROOF. Let M_0 denote the $N \times N$ submatrix of M with $\Delta = \det(M_0)$. By construction,

$$M_0 \begin{bmatrix} \vdots \\ x^\beta \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ x^\beta f_\alpha \\ \vdots \end{bmatrix}, \tag{1.12}$$

where the column vector on the right consists of those $x^\beta f_\alpha$ that give rows of M_0.

(1) Since $\operatorname{Res}(f_0, \dots, f_m)$ is irreducible, it suffices to prove that for any choice of complex coefficients $\overline{c}$,

$$\operatorname{Res}(f_0, \dots, f_m)(\overline{c}) = 0 \Longrightarrow \Delta(\overline{c}) = 0.$$

This is easy, for by (1.11), the vanishing of the resultant for $\overline{c}$ implies that for these coefficients, the system has a nontrivial solution $\overline{x} \in \mathbb{P}^m$. If we evaluate (1.12) at $\overline{c}$ and $\overline{x}$, we get a nonzero element of the kernel of $M_0(\overline{c})$ since $\overline{x} \in \mathbb{P}^m$. This implies that $\Delta(\overline{c}) = \det(M_0(\overline{c})) = 0$.

(2) The classical adjoint M_0^{adj} of M_0 satisfies $M_0^{\text{adj}} M_0 = \Delta I_N$ and has entries in $\mathbb{Q}[c]$. If we multiply each side of (1.12) by M_0^{adj}, we obtain

$$\Delta x^\beta \in \langle f_0, \dots, f_m \rangle$$

for all x^β of degree n. This implies $\Delta \in \langle f_0, \dots, f_m \rangle : \langle x_0, \dots, x_m \rangle^n$, and then (2) follows easily. □

Part (1) of Proposition 1.7 says that when Δ is not identically zero, $\Delta = \operatorname{Res}(f_0, \dots, f_m) E$ for some extraneous factor E. In 1902, Macaulay [**264**] showed how to find $m+1$ carefully chosen maximal minors $D_0, \dots, D_m$ of M so that

$$\operatorname{Res}(f_0, \dots, f_m) = \gcd(D_0, \dots, D_m).$$

He also proved that for the D_i, the extraneous factor E_i is the determinant of an appropriately sized minor of D_i. The gives the formula

$$\operatorname{Res}(f_0, \dots, f_m) = \frac{D_i}{E_i}$$

that expresses the resultant as a quotient of two determinants. Thus we have two formulas for the resultant, both lovely and both wildly impractical for computations due to the large size of the matrices involved.

Part (2) of the proposition is also important. Netto worked affinely, so his version of part (2) can be restated as

$$\Delta \in \mathbb{Z}[c] \cap \langle f_0, \dots, f_m \rangle.$$

Thus Δ is in the elimination ideal $\mathbb{Z}[c] \cap \langle f_0, \dots, f_m \rangle$ in the affine setting. Netto uses this to prove part (1).

When we switch to homogeneous polynomials, the usual elimination ideal is not helpful, since

$$\mathbb{Z}[c] \cap \langle f_0, \dots, f_m \rangle = \{0\}$$

when $f_0, \dots, f_m$ are homogenous of positive degree in $x_0, \dots, x_m$. This is where the *projective elimination ideal* $\mathbb{Z}[c] \cap \big(\langle f_0, \dots, f_m \rangle : \langle x_0, \dots, x_m \rangle^\infty\big)$ enters the picture. A key result is that the resultant generates the projective elimination ideal of $f_0, \dots, f_m$, i.e.,

$$\langle \operatorname{Res}(f_0, \dots, f_m) \rangle = \mathbb{Z}[c] \cap \big(\langle f_0, \dots, f_m \rangle : \langle x_0, \dots, x_m \rangle^\infty\big). \tag{1.13}$$

For a modern proof, see [**226**, Proposition 2.3 and (4.1.2)], and for an introduction to projective elimination ideals, see [**104**, Chapter 8, §5]. Once we know (1.13), part (1) of Proposition 1.7 follows immediately from part (2).

We will encounter projective elimination ideals in Section 1.2 when we discuss van der Waerden's work.

Final Comments. Algebraic geometry in the 19th century was a vast enterprise. The 1892 review article by Brill and Noether [**45**] is over 450 pages long. There was a lot going on besides elimination theory.

One important topic we have not mentioned is enumerative geometry. Many famous results in the field—27 lines on a smooth cubic surface, 3264 conics simultaneously tangent to five general plane conics—date from the 19th century. In one sense, the classic version of Bézout's Theorem is a result in enumerative geometry, since it counts (with multiplicity) the number of intersections of n hypersurfaces in $\mathbb{P}^n$ when the number is finite. We saw this in Example 1.1, where Newton counted tangent lines using Bézout's Theorem. But researchers like Schubert, Chasles and

others developed powerful geometric methods that had nothing to do with elimination. Writing in 1905, Brill comments that elimination theory is

> still outperformed by the well-known geometric counting methods that in the middle of the sixties [the 1860s] and ever since have flooded geometry with its enumerative results; I mean the characteristic method, the correspondence principle, and more recently the principle of conservation of number. [**44**, p. 278]

The last of these, the *principle of conservation of number*, is important to our story.

The "Princip von der Erhaltung der Anzahl" plays a central role in Schubert's 1879 book *Kalkül der abzählende Geometrie*. He explains the principle as follows:

> In algebraic interpretation, it says nothing else than that changing the constants of an equation either leaves the number of their roots unaffected or causes an infinite number of roots by making the equation an identity. [**310**, p. 12]

Though powerful, the principle of conservation of number lacked a rigorous foundation. In fact, at the 1900 International Congress of Mathematicians, Hilbert [**207**] posed the following problem:

> 15. Rigorous foundation of Schubert's enumerative calculus
>
> The problem is *to establish rigorously and with an exact determination of the limits of their validity those geometrical numbers which Schubert especially has determined on the basis of the so-called principle of special position, or conservation of number, by means of the enumerative calculus developed by him.*

Note that Hilbert refers to the principle of conservation of number. But what he says next is even more interesting:

> Although the algebra of today guarantees, in principle, the possibility of carrying out the processes of elimination, yet for the proof of the theorems of enumerative geometry decidedly more is required, namely, the actual carrying out of the process of elimination in the case of equations of special form in such a way that the degree of the final equations and the number of their solutions may be known in advance.

Hilbert is aware of the potential of elimination theory, but like Brill, he recognizes its limitations when applied to enumerative problems.

We will soon see that in the 20th century, the interaction between elimination theory and the principle of conservation of number led to some unexpected developments.

1.2. Elimination Theory in the 20th Century

After reaching a high point in the early 20th century, resultants and elimination theory went through some rough times in the middle part of the century, only to re-emerge stronger than ever in recent times.

A seminal event was the 1890 publication of Hilbert's great paper *Ueber die Theorie der algebraischen Formen* [**205**] that can lay claim to being the first paper in commutative algebra. Besides introducing the Hilbert Basis Theorem, free resolutions, and Hilbert polynomials, Hilbert's proofs also had a non-constructive aspect that was new to algebra. This was followed in 1893 with the Nullstellensatz in [**206**]. The ideas planted by Hilbert led directly to the work of Emmy Noether in the 1920s that revolutionized algebra.

In 1924, the young Bartel van der Waerden spent seven months in Göttingen, studying topology with Kneser and algebra with Noether. His first paper on algebraic geometry [**350**], published in 1926, begins as follows:

> The rigorous foundation of the theory of algebraic varieties ... can be formulated more simply than it has been done so far, without the help of elimination theory, on the sole basis of field theory and of the general theory of ideals in ring domains.

The phrase "without the help of elimination theory" tells us that something fundamental has changed. It is likely that this sentence was actually written by Noether herself; in [**306**], we learn that she often wrote the introductions to papers by beginning students. But as we will see, this phrase represented an aspiration, since elimination continued to be an essential tool for many years.

The Fundamental Theorem of Elimination Theory. This result was stated informally by Kronecker in 1882 and proved by Mertens in 1899 using resultants. In 1926, van der Waerden found a way to avoid resultants by using the Hilbert Basis Theorem and the Nullstellensatz.

We present the theorem in modern notation. Assume that k is algebraically closed and let $f_1, \dots, f_r \in k[t_1, \dots, t_m, x_0, \dots, x_n]$ be homogeneous in $x_0, \dots, x_n$. We think of $t_1, \dots, t_m$ as parameters, so that

$$f_1(t_1, \dots, t_m, x_0, \dots, x_n) = \cdots = f_r(t_1, \dots, t_m, x_0, \dots, x_n) = 0 \tag{1.14}$$

is a system of homogeneous equations in $x_0, \dots, x_n$ whose coefficients depend on the parameters $t_1, \dots, t_m$. The are two questions to ask about (1.14):

- (Geometric) For which choices of parameters does the system (1.14) have a solution in $\mathbb{P}^n$?
- (Algebraic) What constraints do the equations (1.14) impose among the parameters?

The link between the two questions is contained the following result:

THEOREM 1.8 (Fundamental Theorem of Elimination Theory). *In the above situation, there exist $D_1, \dots, D_h \in k[t_1, \dots, t_m]$ such that*

(1) *Given a point $(a_1, \dots, a_m)$ in affine space $\mathbb{A}^m$, the system*

$$f_1(a_1, \dots, a_m, x_0, \dots, x_n) = \cdots = f_r(a_1, \dots, a_m, x_0, \dots, x_n) = 0$$

has a solution in $\mathbb{P}^n \iff D_1(a_1, \dots, a_m) = \cdots = D_h(a_1, \dots, a_m) = 0$.

(2) *Let $W = \mathbf{V}(f_1, \dots, f_r) \subseteq \mathbb{A}^m \times \mathbb{P}^n$. Then in the commutative diagram*

$$\begin{array}{ccc} W & \hookrightarrow & \mathbb{A}^m \times \mathbb{P}^n \\ & \searrow & \downarrow \pi_1 \\ & & \mathbb{A}^m \end{array}$$

where π_1 is projection onto the first factor, we have

$$\pi_1(W) = \mathbf{V}(D_1, \dots, D_h) \subseteq \mathbb{A}^m.$$

(3) *$D_1, \dots, D_h$ can be chosen to be generators of the projective elimination ideal $k[t_1, \dots, t_m] \cap \big(\langle f_1, \dots, f_s \rangle : \langle x_0, \dots, x_n \rangle^\infty\big)$.*

In terms of the above questions, the vanishing of $D_1, \dots, D_h$ tells us which parameters give a system (1.14) with a solution in $\mathbb{P}^n$, and the D_i also tell us the algebraic constraints that the system imposes on the parameters. Note that a

solution in $\mathbb{P}^n$ is nontrivial (the x_i can't all vanish), which is why we use the ideal quotient $\langle f_1,\dots,f_s\rangle : \langle x_0,\dots,x_m\rangle^\infty$ rather than the elimination ideal $\langle f_1,\dots,f_s\rangle$.

The polynomials $D_1,\dots,D_h$ are sometimes called a *resultant system* since they generalize the classical resultant, which by (1.11) defines $\phi(W)$ in Theorem 1.8 when the number of polynomials equals the number of variables.

PROOF OF THEOREM 1.8. First observe that (1) and (2) are clearly equivalent. We will give three proofs to illustrate how ideas about elimination have evolved.

First Proof. In 1899, Mertens gave a proof using resultants [**272**]. We will explain how his argument begins. For $f_1,\dots,f_r$, set $s = \max(\deg f_i)$ and consider all products $x^\alpha f_i$ of degree s. Then set

$$g_j = \sum_{(\alpha,i)} u_{(\alpha,i),j}\, x^\alpha f_i$$

for $j = 0,\dots,n$ and variables $u = (u_{(\alpha,i),j})$. Mertens denotes the resultant of $g_0,\dots,g_n$ with respect to $x_0,\dots,x_n$ by $\operatorname{Res}\binom{g_0\,\dots\,g_n}{x_0\,\dots\,x_n}$. This is a polynomial in the new variables $u_{(\alpha,i),j}$, so we can expand the resultant as a sum

$$\operatorname{Res}\begin{pmatrix} g_0\dots g_n \\ x_0\dots x_n\end{pmatrix} = \sum_\beta D_\beta(t_1,\dots,t_m)u^\beta.$$

This is sometime called the *Kronecker u-resultant*. The resultant system constructed by Mertens starts with the polynomials $D_\beta(t_1,\dots,t_m)$, but more need to be added since the D_β may vanish identically. Mertens knew that the D_β lie in the projective elimination ideal of $f_1,\dots,f_r$, though he didn't say it this way.

Second Proof. We will follow van der Waerden's 1926 proof from [**351**], with a few changes. For simplicity, set $t = (t_1,\dots,t_m)$ and $x = (x_0,\dots,x_n)$, and let $\hat{I}$ denote the projective elimination ideal from part (3) of the theorem. We leave the following two observations to the reader:

(A) Any $g \in \hat{I}$ vanishes on $\pi_1(W) \subseteq \mathbb{A}^m$. Thus $\pi_1(W) \subseteq \mathbf{V}(\hat{I})$.
(B) To prove that $\pi_1(W) = \mathbf{V}(\hat{I})$, it suffices to show that for any $a \in \mathbb{A}^m \setminus \pi_1(W)$, there is $g \in \hat{I}$ such that $g(a) \neq 0$.

Now we give van der Waerden's proof. Like Mertens, he considers all products $x^\alpha f_i$ of degree s, except that now s is an arbitrary positive integer, and like Netto's matrix in Proposition 1.7, van der Waerden gets a matrix $M_s(t)$ with entries in $k[t]$ satisfying

$$M_s(t)\begin{bmatrix}\vdots\\ x^\beta\\ \vdots\end{bmatrix} = \begin{bmatrix}\vdots\\ x^\alpha f_i\\ \vdots\end{bmatrix}. \tag{1.15}$$

He calls this the *dialytic matrix*, a term due to Sylvester. (These days, the name *Macaulay matrix* is more common.) It has $\omega_s = \binom{s+n}{n}$ columns, and let the $\omega_s \times \omega_s$ minors of $M_s(t)$ be

$$D_s^{(1)}(t),\dots,D_s^{(\ell_s)}(t) \in k[t_1,\dots,t_m]. \tag{1.16}$$

If no such minors exist, use 0 instead. An easy modification of the proof of Proposition 1.7 shows that $D_s^{(j)}(t) \in \hat{I}$ for all j.

Following (B), take $a = (a_1, \dots, a_m) \in \mathbb{A}^m \setminus \pi_1(W)$. Then the system

$$f_1(a, x) = \cdots = f_r(a, x) = 0$$

has no solution in $\mathbb{P}^n$. By the Nullstellensatz, $\langle x_0, \dots, x_n \rangle^s \subseteq \langle f_1(a, x), \dots, f_r(a, x) \rangle$ for some s, which implies that when we evaluate (1.15) at $t = a$, every x^β is a linear combination of the $x^\alpha f_i(a, x)$. It follows that $M_s(a)$ has rank ω_s, so that one of the minors (1.16) is nonvanishing at a. Since the minors lie in $\hat{I}$, we have a complete proof that $\pi_1(W) = \mathbf{V}(\hat{I})$.

The final step in the proof is to use the Hilbert Basis Theorem to find finitely many generators $D_1, \dots, D_h$ of $\hat{I}$. This is the desired resultant system.

Third Proof. This proof uses scheme theory, created by Grothendieck starting in 1956 and published with Dieudonné in *Éléments de géométrie algébrique* [**181**] between 1960 and 1967. We will follow Hartshorne's *Algebraic Geometry* [**190**], since many students learn the subject from this masterful exposition. The key result we need is the following.

THEOREM 1.9. *$\mathbb{P}^n_{\mathbb{Z}} \to \operatorname{Spec}(\mathbb{Z})$ is a proper morphism of schemes.*

This is Theorem II.4.9 in [**190**]. Here, $\mathbb{P}^n_{\mathbb{Z}} = \operatorname{Proj}(\mathbb{Z}[x_0, \dots, x_n])$ is projective space over the integers. Also recall that a morphism of schemes $f : X \to Y$ is *proper* if:

- f is *separated* and *of finite type* (see [**190**, Chapter II]).
- f is *universally closed*, meaning that for every morphism $Y' \to Y$, the morphism $X \times_Y Y' \to Y'$ is closed, i.e., maps closed subsets to closed subsets in the Zariski topology.

The map $X \times_Y Y' \to Y'$ is part of a *Cartesian square*

$$\begin{array}{ccc} X \times_Y Y' & \longrightarrow & X \\ \downarrow & \square & \downarrow \\ Y' & \longrightarrow & Y \end{array}$$

The proof of Theorem 1.9 uses the *valuative criterion* that will be explained later in the section.

We now prove part (2) of Theorem 1.8, which says that in the Zariski topology, $\pi_1(W)$ is closed in $\mathbb{A}^m$. The latter is really $\mathbb{A}^m_k = \mathbb{A}^m_{\mathbb{Z}} \times_{\mathbb{Z}} \operatorname{Spec}(k)$, and similarly $\mathbb{P}^m_k = \mathbb{P}^m_{\mathbb{Z}} \times_{\mathbb{Z}} \operatorname{Spec}(k)$. Since $W = \mathbf{V}(f_1, \dots, f_r) \subseteq \mathbb{A}^m_k \times_k \mathbb{P}^m_k$ is closed, the proof is encapsulated in a single commutative diagram of schemes with a Cartesian square:

$$\begin{array}{ccccc} W & \xrightarrow{\text{closed}} & \mathbb{A}^m_k \times_k \mathbb{P}^n_k & \longrightarrow & \mathbb{P}^n_{\mathbb{Z}} \\ & \searrow^{\text{closed image}} & \downarrow \text{closed} & \square & \downarrow \text{proper} \\ & & \mathbb{A}^m_k & \longrightarrow & \operatorname{Spec}(\mathbb{Z}) \end{array}$$

This argument, taken from a 1991 paper of Jouanolou [**226**], reflects Grothendieck's philosophy: once things are set up correctly, basic results become obvious. □

At its heart, Theorem 1.8 is an *existence* theorem. This is reflected in how the three proofs show that $\pi_1(W) = \mathbf{V}(\hat{I})$:

- In Proofs 1 and 2, the key is to create polynomials in $\hat{I}$ that precisely define $\pi_1(W)$. Proof 1 does this using resultants, while Proof 2 replaces resultants with minors of a matrix of polynomials.

- In Proof 3, the valuative criterion guarantees that any point of $\mathbf{V}(\hat{I})$ is a partial solution extends to a solution in W (in the language of [**104**]).

The evolution from Proof 1 to Proof 2 is easy to understand. Van der Waerden knew Mertens's 1899 paper and Macaulay's 1916 book but also understood the full power of what Hilbert did in 1890 and 1893. The introduction to his 1926 paper [**351**] notes that the main result is

> an almost immediate consequence of two fundamental theorems of HILBERT, which may be called HILBERT's Nullstellensats and HILBERT's basis theorem.

Proof 2 is closely related to the proof used in *Ideals, Varieties and Algorithms* [**104**, Theorem 6 of Chapter 8, §5].

The replacement of resultants with the dialytic matrix $M_s(t)$ used in Proof 2 is also interesting. The key step in the proof was to show that if the $f_i(a,x)$ vanish only at $(0,\dots,0)$, then $M_s(a)$ has rank ω_s. Here is how van der Waerden states this result:

> If the H-ideal $(f_1,\dots,f_r)$ has only the solution $(0,\dots,0)$ in Ω, then the dialtyic matrix of the ideal for degree s has rank ω_s, and conversely. [**351**]

(Here, "H-ideal" means homogeneous ideal and Ω is an algebraically closed field.) We will see later that replacing resultants with rank conditions on a matrix of polynomials is a powerful idea that is still used in current research.

Going from Proof 2 to Proof 3 represents a substantial change—something dramatic happened to algebraic geometry between van der Waerden's 1926 paper and Hartshorne's 1977 book. As we will soon see, the Principle of Conservation of Number played a key role in this change.

Sul Principio della Conservazione del Numero. This is the title of a paper written in 1912 by Severi [**318**]. Given the concerns about the Principle of Conservation of Number raised by Hilbert and others, Severi acknowledges that the desire for a precise statement of the principle is

> something more than just a scruple about exaggerated rigor.

In 1971, van der Waerden [**353**] summarized Severi's version of the principle as follows. We have irreducible projective varieties U, V over the complex numbers $\mathbb{C}$ and an irreducible subvariety $W \subseteq U \times V$ such that for a generic $\xi \in U$, there are exactly h points $(\xi,\eta_1),\dots,(\xi,\eta_h) \in W$ lying above ξ. This says that we have diagram

$$\begin{array}{ccc} & W \subseteq U \times V & \\ {\scriptstyle \pi_1}\swarrow & & \searrow{\scriptstyle \pi_2} \\ \xi \in U & & V \ni \eta_1,\dots,\eta_h \end{array} \tag{1.17}$$

where the map π_1 has generic degree h. Take a smooth point $x \in U$ and let ξ approach x, where Severi is thinking topologically. We write this informally as $\xi \rightsquigarrow x$.

Now suppose that $\pi_1^{-1}(x)$ is finite. Then Severi claims that $\xi \rightsquigarrow x$ extends to $(\xi,\eta_i) \rightsquigarrow (x,y_i) \in W$ for $i = 1,\dots,h$, where the points $y_1,\dots,y_h \in V$ have the following properties:

(1) No matter how ξ approaches x, we always get the same points $y_1,\dots,y_h \in V$ up to order.
(2) The obvious inclusion $\{y_1,\dots,y_h\} \subseteq \pi_1^{-1}(x)$ is an equality.

(3) If μ of the η_i have the same limit $y \in V$, we say that y has *multiplicity* μ. Then the sum of the multiplicities μ is equal to h.

This is Severi's version of the Principle of Conservation of Number.

To make this rigorous, we need careful definitions of generic point, multiplicity, and what $\xi \rightsquigarrow x$ really means. We also need to address *extension*: why does $\xi \rightsquigarrow x$ extend to $(\xi, \eta_i) \rightsquigarrow (x, y_i)$? Van der Waerden tackled all of these questions in a series of papers. We will focus on generic points and $\xi \rightsquigarrow x$ but ignore the more substantial topic of multiplicity (interested readers can consult [**287**, **306**, **353**]).

Generic Points. In their 1915 book on algebraic geometry, Enriques and Chisini define a generic point of an irreducible variety as follows:

> The notion of a generic 'point' or 'element' of a variety, i.e., the distinction between properties that pertain in general to the points of a variety and properties that only pertain to exceptional points, now takes on a precise meaning for all algebraic varieties.
>
> A property is said to pertain in general to the points of a variety V_n, of dimension n, if the points of V_n not satisfying it form – inside V_n – a variety of less than n dimensions.

In their view, "generic" is a relative concept that depends on the property in question. A point can be generic for one property but not generic for another. This is still one of the standard meanings of "generic."

In his first 1926 paper [**350**], van der Waerden proposes an absolute notion of "generic," where a generic point lives in an extension of our original field k. For simplicity, we assume that k is algebraically closed. Let

$$\mathfrak{p} \subseteq R = k[x_1, \ldots, x_n]$$

be a prime ideal. This gives the quotient ring $R/\mathfrak{p}$ with field of fractions Ω. Note that $k \subseteq \Omega$. Geometrically, $\mathfrak{p}$ defines an affine variety V, and for any extension field $k \subseteq L$, we write $V(L)$ to denote the solutions over L, i.e.,

$$V(L) = \{(a_1, \ldots, a_n) \in L^n \mid f(a_1, \ldots, a_n) = 0 \text{ for all } f \in \mathfrak{p}\}.$$

Van der Waerden constructs a particular point $\xi = (\xi_1, \ldots, \xi_n) \in V(\Omega)$ as follows. For each i, consider the coset $\xi_i = x_i + \mathfrak{p} \in R/\mathfrak{p} \subseteq \Omega$. Then one sees immediately that if $f \in k[x_1, \ldots, x_n]$, then

$$f(\xi_1, \ldots, \xi_n) = 0 \iff f(x_1, \ldots, x_n) + \mathfrak{p} = 0 + \mathfrak{p} \iff f \in \mathfrak{p}. \tag{1.18}$$

This shows that $\xi = (\xi_1, \ldots, \xi_n) \in V(\Omega)$. But this is not just any point of $V(\Omega)$, for by (1.18), any $f \in k[x_1, \ldots, x_n]$ that vanishes at the single point $\xi \in V(\Omega)$ also vanishes at *every* point of $V(L)$ for *every* $k \subseteq L$. For this reason, van der Waerden calls ξ an "allgemeine Nullstelle" ("general zero") of $\mathfrak{p}$. From this point of view, a property is generic for V if and only if it holds at ξ.

This notion of generic point is similar to the generic points defined in Weil's 1946 book [**359**, Chapter IV, §1]. It is also related to what happens in scheme theory, where the prime ideal $\mathfrak{p} \subseteq R = k[x_1, \ldots, x_n]$ gives the point

$$\mathfrak{p} \in \mathbf{V}(\mathfrak{p}) \subseteq \mathbb{A}^n_k = \operatorname{Spec}(R).$$

In this situation, a polynomial $f \in R$ has the value $f(\mathfrak{p}) \in k(\mathfrak{p})$ given by the image of f in the residue field of $\mathcal{O}_{\mathbb{A}^n_k, \mathfrak{p}}$, which is simply van der Waerden's field Ω in

fancier language. It follows that

$$f(\mathfrak{p}) = 0 \iff f \in \mathfrak{p}. \tag{1.19}$$

This is the modern version of (1.18).

Specialization. For Severi, $\xi \rightsquigarrow x$ used the usual topology of $\mathbb{C}$. In 1927, van der Waerden [**352**] figured out how to do this algebraically using what he called a "relationstreue Spezialisierung" ("relation-preserving specialization"). Later this was shortened to *specialization*.

We explore a simple case of specialization that focuses on diagram (1.17) and uses van der Waerden's generic point $\xi \in U(\Omega)$. Let $\overline{\Omega}$ be the algebraic closure of Ω and suppose that we have $\eta \in V(\overline{\Omega})$ such that $(\xi, \eta) \in W(\overline{\Omega})$. Recall that U and V are projective, so that $U \times V \subseteq \mathbb{P}^m \times \mathbb{P}^n$ for some m, n.

In this situation, we say that a point $(x, y) \in W(k)$ is a *specialization* (in the sense of van der Waerden) of (ξ, η), written $(\xi, \eta) \rightsquigarrow (x, y)$, provided that

$$H(\xi, \eta) = 0 \Longrightarrow H(x, y) = 0 \tag{1.20}$$

for all bihomogeneous polynomials H in the coordinate ring $k[\mathbb{P}^m \times \mathbb{P}^n]$. Then we have the following existence result.

PROPOSITION 1.10. *With the above setup, take any point $x \in U(k)$. Then there exists a point $y \in V(k)$ such that $(x, y) \in W(k)$ and $(\xi, \eta) \rightsquigarrow (x, y)$.*

PROOF. The ideal of all $H \in k[\mathbb{P}^m \times \mathbb{P}^n]$ that satisfy $H(\xi, \eta) = 0$ defines a subvariety $W_0 \subseteq U \times V$ such that $(\xi, \eta) \in W_0(\overline{\Omega})$. We have the diagram

$$\begin{array}{ccc} W_0 & \hookrightarrow & U \times V \\ & \searrow & \downarrow \pi_1 \\ & & U \end{array}$$

where V is projective. With a little thought, one sees that Theorem 1.8 implies that $\pi_1(W_0) \subseteq U$ is closed. Since $\pi_1(\xi, \eta) = \xi$ is a generic point of U, it follows that $\pi_1(W_0) = U$, so that $x \in U$ lifts to $(x, y) \in W_0$. Then (1.20) follows immediately from the definition of W_0. □

Note also that since ξ is a generic point of U, we have a specialization $\xi \rightsquigarrow x$ if and only if $x \in U$—this follows immediately from the projective version of (1.18). Thus we can restate Proposition 1.10 as a result about extending specializations.

PROPOSITION 1.11 (Extension of Specializations). *Given (ξ, η) as above, any specialization $\xi \rightsquigarrow x$ extends to a specialization $(\xi, \eta) \rightsquigarrow (x, y)$.*

Van der Waerden was aware that the proof of Proposition 1.10/1.11 required the Fundamental Theorem of Elimination Theory. So the goal stated in his 1926 paper [**350**] of finding a purely algebraic foundation for algebraic geometry "without the help of elimination theory" remained unfulfilled. Nevertheless, after tackling multiplicities in his 1927 paper [**352**], van der Waerden was able to provide a complete proof of Severi's version of the Principle of Conservation of Number.

We should also say a few words about what specialization means in scheme theory. For concreteness, let $R = k[x_1, \ldots, x_n]$, k algebraically closed, so that

maximal ideals of R correspond to points in k^n. These are the closed points of the affine scheme $\mathbb{A}^n_k = \operatorname{Spec}(R)$. Given points $\mathfrak{p}, \mathfrak{q} \in \mathbb{A}^n_k$, we have the equivalences

$$\begin{aligned} \mathfrak{q} \in \overline{\{\mathfrak{p}\}} &\iff \mathfrak{p} \subseteq \mathfrak{q} \\ &\iff \text{for all } f \in R,\ f(\mathfrak{p}) = 0 \Longrightarrow f(\mathfrak{q}) = 0, \end{aligned}$$

where we are using the Zariski topology on $\mathbb{A}^n_k$ and $f(\mathfrak{p}) = 0$ has the same meaning it did in (1.19). In this situation, we say that $\mathfrak{q}$ is a *specialization* of $\mathfrak{p}$ and write $\mathfrak{p} \rightsquigarrow \mathfrak{q}$. The relation to (1.20) is clear.

Weil and Elimination. In 1946, Weil published *Foundations of Algebraic Geometry* [**359**], where he drew on ideas of van der Waerden, Zariski and others to create his own unique approach to algebraic geometry. As already mentioned, Weil defines a version of "generic point" similar to van der Waerden's, and the same is true for the definition of specialization given in [**359**, Chapter II, §1].

In his book, the title of Chapter II, §2 is "Extension of Specializations." This is where Weil proves his version of Proposition 1.11. The crucial difference is that *he does not use elimination theory*. Rather, he proves that specializations extend by means of a "device" of Chevalley. In a footnote on page 31, Weil says that Chevalley's device,

> it may be hoped, finally eliminates from algebraic geometry the last traces of elimination-theory

This is the only mention of elimination theory in Weil's book. In his view, it has no role in algebraic geometry.

The Valuative Criterion. In Chapter II of [**359**], Weil was trying to find the essence of what it means to extend a specialization. This process continued with Grothendieck, who realized that Zariski's theory of *valuation rings* provides the right tool to understand the heart of the matter. This led to the *valuative criterion*, which we now explain.

An integral domain R with field of fractions K is a *valuation ring* if for every $\alpha \in K$, we have either $\alpha \in R$ or $\alpha^{-1} \in R$. Then R is a local ring with maximal ideal $\mathfrak{m} = \{\alpha \in R \mid \alpha^{-1} \notin R\}$. The ring R has distinguished prime ideals $\{0\}$ and $\mathfrak{m}$, which are the *generic* and *special* points of the affine scheme $\operatorname{Spec}(R)$. The inclusion $R \subseteq K$ gives a morphism $\operatorname{Spec}(K) \to \operatorname{Spec}(R)$ that sends the unique point of $\operatorname{Spec}(K)$ to the generic point of $\operatorname{Spec}(R)$. A comprehensive account of valuation rings appears in [**368**, Chapter VI].

Now suppose that Y is a scheme with points $\xi, y \in Y$. Then one can prove that $\xi \rightsquigarrow y$ if and only if there exists a valuation ring R and a morphism $\operatorname{Spec}(R) \to Y$ that sends the generic and special points of $\operatorname{Spec}(R)$ to ξ and y respectively.

This allows us to formulate an especially nice case of what it means to extend a specialization $\xi \rightsquigarrow y$ in Y. Suppose that we have a commutative diagram

$$\begin{array}{ccc} \operatorname{Spec}(K) & \longrightarrow & X \ni \xi' \\ \downarrow & \nearrow & \downarrow \\ \operatorname{Spec}(R) & \longrightarrow & Y \ni \xi, y \end{array} \tag{1.21}$$

where $\operatorname{Spec}(R) \to Y$ is as above and $\xi' \in X$ maps to $\xi \in Y$. When the dotted arrow exists in (1.21) making the whole diagram commutative, the special point of $\operatorname{Spec}(R)$ maps to $x \in X$ giving a specialization $\xi' \rightsquigarrow x$ in X that extends $\xi \rightsquigarrow y$.

The valuative criterion states that a separated morphism of finite type $X \to Y$ with X Noetherian is proper if and only the dotted arrow exists for every diagram (1.21) where R is a valuation ring. This is proved in [**190**, Theorem II.4.7]. The result is a lovely sequence of ideas:

- A rigorous foundation for the Principle of Conservation of Number requires the ability to extend specializations.
- This was first proved by the Fundamental Theorem of Elimination Theory, which was then eliminated with ideas that led to the valuative criterion for properness.
- The valuative criterion gives a quick proof that $\mathbb{P}^n_{\mathbb{Z}} \to \operatorname{Spec}(\mathbb{Z})$ is proper, which has the Fundamental Theorem of Elimination Theory as an immediate corollary.

Abhyankar and Elimination. Besides wonderful mathematics, there is also a human side to this story. At the 1970 International Congress of Mathematicians, van der Waerden gave a talk entitled *The Foundation of Algebraic Geometry from Severi to André Weil*, where he sketched the developments from 1912 to 1946 that we have been discussing. His lecture was later published as [**353**].

By 1970, the Grothendieck revolution in algebraic geometry was in full swing. One of the people who heard van der Waerden speak at the ICM was Abhyankar, a student of Zariski who had a deep affection for 19th century algebraic geometry and no love for the extreme abstraction of the Grothendieck school. A poem he published in 1971 begins as follows:

> Polynomials and power series
> May they forever rule the world.
>
> Eliminate, eliminate, eliminate
> Eliminate the eliminators of elimination theory.
>
> As you must resist the superbourbaki coup
> So must you fight the little bourbakis too.
>
> Kronecker, Kronecker, Kronecker above all
> Kronecker, Mertens, Macaulay, and Sylvester.
>
> Not the theology of Hilbert
> But the constructions of Gordan. [**3**]

Most of the names are familiar. Abyhankar imagines a conflict between the new and the old, making it clear where his sympathies lie. Fortunately, there is no conflict—scheme theory has proved to be a powerful, flexible language that has enabled algebraic geometry to not only give rigorous proofs of the geometric insights of the 19th century but also give rise to whole new areas of inquiry, including the minimal model program, arithmetic algebraic geometry, and a host of applications.

The Rebirth of Elimination. When Weil's book appeared in 1946, elimination theory was at low ebb. New impetus came with the 1965 PhD thesis of Bruno Buchberger [**47**] and his 1970 article [**48**] that introduced Gröbner bases, the Buchberger criterion and Buchberger algorithm. A few years later, Trinks proved the following result in a paper published in 1978. If $\mathfrak{R} = \mathfrak{r}[X_1, \dots, X_n]$, then:

> KOROLLAR. *Ist* $\leq$ *die lexikographische Ordnung,* $\mathfrak{R}_\nu = \mathfrak{r}[X_1, \dots, X_\nu]$ $(0 \leq \nu \leq n)$ *und* $\mathfrak{G}$ *eine mittels* $\leq$ *gebildte G-Basis von* $\mathfrak{A} \trianglelefteq \mathfrak{R}$*, so ist* $\mathfrak{G}_\nu := \mathfrak{G} \cap \mathfrak{R}_\nu$ *eine G-Basis von* $\mathfrak{A}_\nu := \mathfrak{A} \cap \mathfrak{R}_\nu$*.* [**347**]

This is now known as the Elimination Theorem. Elimination theory is back in business after a 20 year hiatus! Furthermore, by the 1970s, researchers had an algorithmic basis for elimination, independent of resultants, and computers that allowed them to implement these algorithms. This led to the field of computational commutative algebra and algebraic geometry, which continues to be an active area of research today.

Besides these developments, other interesting things happened in the 1970s. Here are a few highlights:

- In 1975, Bernstein's paper *The number of roots of a system of equations* [**25**] proved a version of Bézout's Theorem that used mixed volume.
- In 1975–76, Giraud taught a course on elementary algebraic geometry [**169**] that discussed the Sylvester resultant.
- In 1976, Abhyankar wrote an impassioned defense of elementary algebra (what he called "high-school algebra") in algebraic geometry [**4**].
- In 1977, Lazard published *Algèbre linéaire sur* $K[X_1, \cdots, X_n]$*, et élimination* [**252**] as part of a project begun in 1974 on commutative algebra.
- In 1978, Wu published *On the decision problem and the mechanization of theorem-proving in elementary geometry* [**364**], which drew on the algorithmic tradition in classical Chinese mathematics to create an efficient approach to systems of equations.
- In 1979, Jouanolou wrote *Singularités rationelles du résultant* [**224**], which applied the full power of modern algebraic geometry to the resultant.

It is interesting that Lazard, Giraud and Jouanolou were trained in Grothendieck-style algebraic geometry yet recognized the value of resultants and elimination theory. We will say more about Bernstein's work below.

The Rebirth of Resultants. To get a better sense of the status of resultants in 1979, let's look more closely at Jouanolou's paper [**224**]. He fixes a Noetherian ring k and considers homogenous polynomials $f_0, \dots, f_n$ of degree $d_0, \dots, d_n$ in $X_0, \dots, X_n$ whose coefficients are independent variables. Let $A = k[c]$, where c is the vector of coefficients of $f_0, \dots, f_n$, and set

$$\mathbb{A}_k^N = \operatorname{Spec}(A),\ W = \operatorname{Proj}(A[X_0, \dots, X_n]/\langle f_0, \dots, f_n\rangle).$$

In down-to-earth terms, points of $\mathbb{A}_k^N$ parametrize all possible choices of the coefficients $\overline{c}$ of $f_0, \dots, f_n$, and W consists of all pairs $(\overline{c}, \overline{x}) \in \mathbb{A}_k^N \times \mathbb{P}_k^n$ such that

$$f_0(\overline{c}, \overline{x}) = \cdots = f_n(\overline{c}, \overline{x}) = 0. \tag{1.22}$$

The resultant $\operatorname{Res}_{d_0,\dots,d_n}(f_0, \dots, f_n)$ is a polynomial in $\mathbb{Z}[c]$ that vanishes whenever the equations (1.22) have a solution. It follows that we have a commutative diagram of schemes:

$$\begin{array}{ccc} W & \hookrightarrow & \mathbb{A}_k^N \times \mathbb{P}_k^n \\ \downarrow & & \downarrow \\ \mathbf{V}(\operatorname{Res}_{d_0,\dots,d_n}) & \hookrightarrow & \mathbb{A}_k^N \end{array} \tag{1.23}$$

Jouanolou proves that the map $W \to \mathbf{V}(\operatorname{Res}_{d_0,\dots,d_n})$ is a *resolution of singularities* and uses duality to study where $\mathbf{V}(\operatorname{Res}_{d_0,\dots,d_n})$ has *rational singularities*.

In [**224**], Jounanolou cites Macaulay's 1902 paper and 1916 book since in 1979, these were the best references available for the properties of resultants. So his next

project was to write a series of papers [**225**] (1980), [**226**] (1991), [**227**] (1995), [**228**] (1996) and [**229**] (1997) that not only created the modern theory of the classical resultant but also revealed its intimate connection to commutative algebra. When writing the chapter on resultants in *Using Algebraic Geometry* [**105**], Jouanolou's paper [**226**] from 1991 was an invaluable resource.

One recent contribution is Staglianó's Macaulay2 package *Resultants* [**328**] for computing the classical resultant, published in 2018.

Bernstein's Theorem. Another important development was the theory of *sparse polynomials*. We have seen scattered examples going back to Bézout, but in the 19th century much of the focus was on complete polynomials and their homogeneous analogs. In the sparse case, we consider *Laurent polynomials* $f \in k[x_1^{\pm 1}, \dots, x_n^{\pm 1}]$, where we now work affinely and allow negative exponents. The exponent vectors of f lie in $\mathbb{Z}^n$, and their convex hull is a lattice polytope denoted $N(f) \subseteq \mathbb{R}^n$.

The algebraic geometry of sparse polynomials was studied by Bernstein, Khovanskii and Kouchnirenko in the 1970s. A key result is Bernstein's theorem from 1975 [**25**]. Here is a version due to Fulton [**157**, Section 5.5]:

THEOREM 1.12. *Suppose $f_1, \dots, f_n \in \mathbb{C}[x_1^{\pm 1}, \dots, x_n^{\pm 1}]$ have Newton polytopes $Q_1, \dots, Q_n$, and assume that $\mathbb{Z}^n$ is affinely generated by the exponent vectors of $f_1, \dots, f_n$. Then for*

$$Z = \{x \in (\mathbb{C}^*)^n \mid f_1(x) = \cdots = f_n(x) = 0\},$$

we have

$$\sum_{z \in Z \text{ isolated}} i(z; f_1, \dots, f_n) \leq MV_n(Q_1, \dots, Q_n).$$

Because of the interactions between Bernstein, Khovanskii and Kouchnirenko, Theorem 1.12 is sometimes referred to as the *BKK bound*. A standard refinement of Theorem 1.12 is that if the coefficients of the f_i are generic, then the solution set Z is finite and consists of precisely $MV_n(Q_1, \dots, Q_n)$ points. See, for example, [**165**, Theorem 6.2.8]. Let's say a bit more about Bernstein's theorem:

- We work in $(\mathbb{C}^*)^n$ because $f_1, \dots, f_n$ are Laurent polynomials.
- An isolated point $z \in Z$ has intersection multiplicity

$$i(z; f_1, \dots, f_n) = \dim \mathcal{O}_{(\mathbb{C}^*)^n, z} / \langle f_1, \dots, f_n \rangle.$$

- $MV_n(Q_1, \dots, Q_n)$ is the n-dimensional *mixed volume* of $Q_1, \dots, Q_n$. We recall the definition in Example 1.13 below. Note that we use the version that satisfies $MV_n(Q, \dots, Q) = n! \operatorname{Vol}_n(Q)$. See [**105**, §7.4] for more details.
- Fulton's version of Bernstein's theorem is from his book *Introduction to Toric Varieties* [**157**]. In Section 5.4, he explains the toric connection by constructing a projective toric variety X_Σ such that $f_1, \dots, f_n$ correspond to sections of line bundles $L_1, \dots, L_n$ on X_Σ whose intersection product is

$$L_1 \cdot L_2 \cdots L_n = MV_n(Q_1, \dots, Q_n).$$

  This is Bézout's Theorem for the toric variety X_Σ.

In Section 1.1, we saw some interesting examples of sparse polynomials from the 18th and 19th centuries. Let's return to a case considered by Bézout.

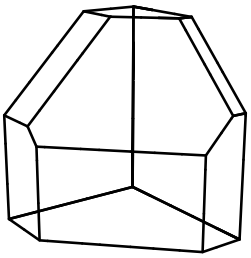
EXAMPLE 1.13. Recall from Example 1.4 that a polynomial f in x, y, z is of the *third species* if there are positive integers $t, a_1, a_2, a_3, b_1, b_2, b_3$ such that the monomials $x^{k_1}y^{k_2}z^{k_3}$ of f satisfy

$$\begin{aligned} &k_1 + k_2 + k_3 \le t \\ &0 \le k_1 \le a_1, \quad 0 \le k_2 \le a_2, \quad 0 \le k_3 \le a_3 \\ &k_1 + k_2 \le b_3, \quad k_1 + k_3 \le b_2, \quad k_2 + k_3 \le b_1. \end{aligned} \tag{1.24}$$

Here, we replace Bézout's awkward $a, \underset{\prime}{a}, \underset{\prime\prime}{a}$ with a_1, a_2, a_3, and similarly for b_1, b_2, b_3. The *first form* of the third species is when $t, a_1, a_2, a_3, b_1, b_2, b_3$ satisfy

$$t - b_1 - b_2 + a_3 < 0, \quad t - b_1 - b_3 + a_2 < 0, \quad t - b_2 - b_3 + a_1 < 0.$$

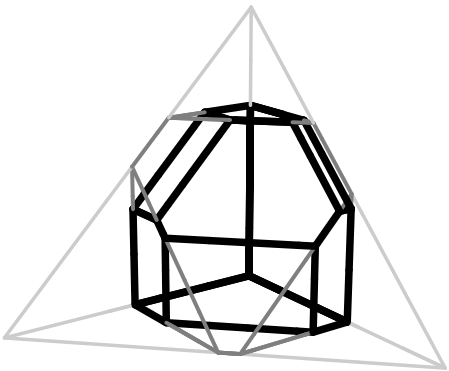
This give a polytope denoted $P = P(t, a_1, a_2, a_3, b_1, b_2, b_3)$. Here is a typical picture:



To compute the Euclidean volume $\mathrm{Vol}_3(P)$ of P, we start with the simplex

$$\Delta_3 = \{(a, b, c) \in \mathbb{R}^3 \mid a, b, c \ge 0, a + b + c \le 1\} \subseteq \mathbb{R}^3,$$

which has volume $\mathrm{Vol}_3(\Delta_3) = 1/6$. The first line of (1.24) tells us that P sits inside $t\Delta_3$. Then the next line of (1.24) chops off three corners of $t\Delta_3$:



The corners, shown in light gray, are translates of $(t - a_1)\Delta_3$, $(t - a_2)\Delta_3$ and $(t - a_3)\Delta_3$. It follows that the removing the corners leaves us with the volume

$$\tfrac{1}{6}t^3 - \tfrac{1}{6}(t - a_1)^3 - \tfrac{1}{6}(t - a_2)^3 - \tfrac{1}{6}(t - a_3)^3.$$

It remains to consider the third line of (1.24), which removes the three prisms shown in darker gray. The reader can verify that the bottom prism has volume

$$\tfrac{1}{6}(t - b_3)^2(3a_1 + 3a_2 - 2b_3 - t) = -\tfrac{1}{6}(t - b_3)^3 + \tfrac{1}{2}(a_1 + a_2 - b_3)(t - b_3)^2.$$

Using similar formulas for the other two prisms, we get

$$\begin{aligned}\mathrm{Vol}_3(P) = \tfrac{1}{6}\big(&t^3-(t-a_1)^3-(t-a_2)^3-(t-a_3)^3+(t-b_1)^3\\ &+(t-b_2)^3+(t-b_2)^3-3(a_1+a_2-b_3)(t-b_3)^2\\ &-3(a_1+a_3-b_2)(t-b_2)^2-3(a_2+a_3-b_1)(t-b_1)^2\big).\end{aligned}\tag{1.25}$$

Now suppose that $f,g,h\in\mathbb{C}[x,y,z]$ have $N(f)=N(g)=N(h)=P$ and f,g,h are generic within all polynomials with this Newton polytope. Then all solutions of $f=g=h=0$ in $\mathbb{C}^3$ lie in $(\mathbb{C}^*)^3$, and the number solutions is given by

$$\begin{aligned}MV_3(P,P,P) &= 3!\,\mathrm{Vol}_3(P)\\ &= t^3-(t-a_1)^3-(t-a_2)^3-(t-a_3)^3+(t-b_1)^3\\ &\quad+(t-b_2)^3+(t-b_2)^3-3(a_1+a_2-b_3)(t-b_3)^2\\ &\quad-3(a_1+a_3-b_2)(t-b_2)^2-3(a_2+a_3-b_1)(t-b_1)^2,\end{aligned}\tag{1.26}$$

where the first equality is a standard property of mixed volume and the second follows from (1.25).

Since (1.26) is the number of solutions, this should be the degree of Bézout's "équation finale résultant," denoted D in Example 1.4. If you take formula for D given in (1.7) and set

$$t=t'=t'',\ a_1=a=a'=a'',\ a_2=\underset{,}{a}=\underset{,}{a}'=\underset{,}{a}'',\ldots,\ b_3=\underset{,,}{b}=\underset{,,}{b}'=\underset{,,}{b}'',\tag{1.27}$$

then D reduces to (1.26) on the nose!

But it gets better, for suppose now that we have f,g,h with

$$\begin{aligned}N(f)&=P=P(t,a_1,a_2,a_3,b_1,b_2,b_3)\\ N(g)&=P'=P(t',a_1',a_2',a_3',b_1',b_2',b_3')\\ N(h)&=P''=P(t'',a_1'',a_2'',a_3'',b_1'',b_2'',b_3''),\end{aligned}$$

all of the first form of the third species. Then we have the following nice result.

PROPOSITION 1.14. *Bézout's formula for D from* (1.7) *satisfies the equation*

$$D=MV_3(P,P',P'').$$

PROOF. The mixed volume $MV_3(P,P',P'')$ is constructed by first expanding the volume of the Minkowski sum

$$\mathrm{Vol}_3(\lambda_1P+\lambda_2P'+\lambda_3P'')\tag{1.28}$$

as a polynomial in $\lambda_1,\lambda_2,\lambda_3$. Then the mixed volume $MV_3(P,P',P'')$ is defined to be the coefficient of $\lambda_1\lambda_2\lambda_3$. If we set $\mathbf{t}=(t,a_1,a_2,a_3,b_1,b_2,b_3)$ and similarly for $\mathbf{t}'$ and $\mathbf{t}''$, then

$$\lambda_1P+\lambda_2P'+\lambda_3P''=P(\lambda_1\mathbf{t}+\lambda_2\mathbf{t}'+\lambda_3\mathbf{t}'').$$

Since P,P',P'' are of the first form, the same is true for the above Minkowski sum. So the volume (1.28) is computed using the *same* polynomial that appears in (1.25). Thus $MV_3(P,P',P'')$ is a polynomial in $\mathbf{t},\mathbf{t}',\mathbf{t}''$, which we write as

$$MV_3(P,P',P'')=MV(\mathbf{t},\mathbf{t}',\mathbf{t}'').$$

One can check that MV has the following properties:

- SYMMETRIC: $MV(\mathbf{t},\mathbf{t}',\mathbf{t}'')=MV(\mathbf{t}',\mathbf{t},\mathbf{t}'')=MV(\mathbf{t}',\mathbf{t}'',\mathbf{t})=\cdots$.
- MULTILINEAR: $MV(\mathbf{t},\mathbf{t}',\mathbf{t}'')$ has degree 1 in each of $\mathbf{t},\mathbf{t}',\mathbf{t}''$.
- DIAGONAL: $MV(\mathbf{t},\mathbf{t},\mathbf{t})=3!\,\mathrm{Vol}_3(P(\mathbf{t}))$. This is the polynomial in (1.26).

This means that $MV(\mathbf{t}, \mathbf{t}', \mathbf{t}'')$ is the *blossom* of the polynomial in (1.26). We recommend [**176**] for a nice introduction to blossoming.

But now look at Bézout's formula for D from (1.7). It has the same symmetric and multilinear properties, and in (1.27), we showed that on the diagonal, it reduces to $3!\,\mathrm{Vol}_3(P(\mathbf{t}))$. So D is also a blossom of (1.26). Since the symmetric, multilinear and diagonal properties uniquely characterize the blossom, the result follows. □

Bézout did not think geometrically about polynomials of the first form of the third species, so his derivation of D (described in [**288**]) uses totally different methods. But it is breathtaking to see such a wonderful example of mixed volume in a book written in 1779. ◁▷

Sparse Resultants. Just as Bézout's theorem for complete polynomials led to the classical resultant, it should be no surprise that Bézout's theorem for sparse polynomials led to the sparse resultant. Here is the basic idea. For concreteness, we will work over the complex numbers. Suppose we have finite sets $\mathcal{A}_0, \ldots, \mathcal{A}_n \subseteq \mathbb{Z}^n$. The resulting *Laurent polynomials*

$$f_i(x_1, \ldots, x_n) = \sum_{a \in \mathcal{A}_i} c_{i,a}\, x^a, \quad i = 0, \ldots, n \tag{1.29}$$

give a system of $n+1$ equations $f_0 = \cdots = f_n = 0$ in $(\mathbb{C}^*)^n$. Since this space has dimension n, the system usually has no solutions. But for special choices of the coefficients $c_{i,a}$, solutions can exist. This is what the sparse resultant does.

For the rest of our discussion of sparse resultants, we will make two simplifying assumptions:

- The convex hulls $Q_i = \mathrm{Conv}(\mathcal{A}_i) \subseteq \mathbb{R}^n$ have dimension n for $i = 0, \ldots, n$.
- $\mathbb{Z}^n$ is affinely generated by $\mathcal{A}_0 \cup \cdots \cup \mathcal{A}_n$,

The coefficients $c_{i,a}$ of $f_0, \ldots, f_n$ give a point in $\mathbb{A}^N$, where $N = \sum_{i=0}^n |\mathcal{A}_i|$ is the total number of coefficients. Set

$$W = \{(c_{i,a}, x) \in \mathbb{A}^N \times (\mathbb{C}^*)^n \mid f_0(x) = \cdots = f_n(x) = 0\} \subseteq \mathbb{A}^N \times (\mathbb{C}^*)^n,$$

and let $Z \subseteq \mathbb{A}^N$ is the Zariski closure of the projection of W. Here is the basic result [**165**, 8.1].

THEOREM 1.15. *With the above hypotheses, there is an absolutely irreducible polynomial*

$$\mathrm{Res}_{\mathcal{A}_0,\ldots,\mathcal{A}_n} \in \mathbb{Z}[c_{i,a} \mid 0 \le i \le n,\, a \in \mathcal{A}_i]$$

with relatively prime coefficients such that $Z = \mathbf{V}(\mathrm{Res}_{\mathcal{A}_0,\ldots,\mathcal{A}_n})$. *Furthermore, for a fixed* i, *the polynomial* $\mathrm{Res}_{\mathcal{A}_0,\ldots,\mathcal{A}_n}$ *is homogeneous in the coefficients* $c_{i,a}$ *of* f_i *of degree* $MV_n(Q_0, \ldots, \widehat{Q_i}, \ldots, Q_n)$.

It follows that we have a commutative diagram

$$\begin{array}{ccc} W & \hookrightarrow & \mathbb{A}^N \times (\mathbb{C}^*)^n \\ \downarrow & & \downarrow \\ \mathbf{V}(\mathrm{Res}_{\mathcal{A}_0,\ldots,\mathcal{A}_n}) & \hookrightarrow & \mathbb{A}^N \end{array} \tag{1.30}$$

$\mathrm{Res}_{\mathcal{A}_0,\ldots,\mathcal{A}_n}$ is sometimes called the *mixed sparse resultant* (since each polynomial is allowed to have different support).

EXAMPLE 1.16. The classical resultant $\mathrm{Res}_{d_0,\dots,d_n}$ is the special case of the sparse resultant where $\mathcal{A}_i = (d_i\Delta_n) \cap \mathbb{Z}^n$ and Δ_n is the standard n-simplex:

$$\Delta_n = \{(a_1,\dots,a_n) \in \mathbb{R}^n \mid a_1,\dots,a_n \geq 0, a_1 + \cdots + a_n \leq 1\}.$$

Then (1.30) is a version of diagram (1.23) that uses $(\mathbb{C}^*)^n$ in place of $\mathbb{P}^n$. ◁▷

EXAMPLE 1.17. Let $\mathcal{A} = \mathcal{A}_1 = \cdots = \mathcal{A}_n \subseteq \mathbb{Z}^n$, where $Q = \mathrm{Conv}(\mathcal{A}) \subseteq \mathbb{R}^n$ has dimension n and $\mathbb{Z}^n$ is affinely generated by $\mathcal{A}$. Then $\mathrm{Res}_{\mathcal{A}_0,\dots,\mathcal{A}_n}(f_0,\dots,f_n)$ is often written $\mathrm{Res}_{\mathcal{A}}(f_0,\dots,f_n)$, called the *sparse $\mathcal{A}$-resultant* or *unmixed sparse resultant* (since all polynomials have the same support).

The lattice polytope Q gives a toric variety X_Q with torus $(\mathbb{C}^*)^n$, and $f_0,\dots,f_n$ can be homogenized with respect to the total coordinate ring of X_Q to give polynomials $F_0,\dots,F_n$ (see [**98**, Section 13]). Then the equations $F_0 = \cdots = F_n = 0$ define $W \subseteq \mathbb{A}^N \times X_Q$, where as above the coefficients of the F_i live in $\mathbb{A}^N$. In this situation, the diagram (1.30) becomes

$$\begin{array}{ccc} W & \hookrightarrow & \mathbb{A}^N \times X_Q \\ \downarrow & & \downarrow \\ \mathbf{V}(\mathrm{Res}_{\mathcal{A}}) & \hookrightarrow & \mathbb{A}^N \end{array}$$

In (1.30), the resultant vanishes when the equations have a solution in $(\mathbb{C}^*)^n$, while in this diagram, the resultant vanishes *if and only if* the equations have a solution in X_Q. There is a deep link between sparse resultants and toric varieties. ◁▷

Sparse resultants are discussed in detail in *Discriminants, Resultants, and Multidimensional Determinants* [**165**], published in 1994 by Gel'fand, Kapranov and Zelevinski. This astonishing book contains a wealth of material, including the important concept of *discriminant* that we have completely ignored.

We have now arrived at the end of the 20th century, with many omissions. But what about applications? We will give one example that describes what happened when geometric modeling first became aware of elimination theory.

Geometric Modeling. In the early 1980s, a PhD student in mechanical engineering named Tom Sederberg read the book *Computational Geometry for Design and Manufacture*, published in 1979. When discussing the intersection of parametrized surfaces in $\mathbb{R}^3$, the authors write:

> Great advantage would be had if one surface could be translated into implicit form. Unfortunately, the most commonly used parametric surfaces, the bicubic patches, cannot be expressed this way. [**146**, p. 258]

Parametrized surfaces will be covered in Chapter 3, so here, we will say just enough to understand the quote. A *bicubic patch* in $\mathbb{R}^3$ is a surface parametrization

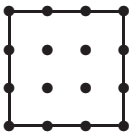
$$x = \sum_{i,j=0}^{3} a_{ij}\, s^i t^j, \quad y = \sum_{i,j=0}^{3} b_{ij}\, s^i t^j, \quad z = \sum_{i,j=0}^{3} c_{ij}\, s^i t^j, \tag{1.31}$$

and its *implicit form* is the equation $F(x,y,z) = 0$ of the surface. This is clearly a problem in elimination theory, which according to [**146**] can't be done! As the first book on geometric modeling, [**146**] reflected what the newly emerging field of geometric modeling knew about algebra at the time.

Sederberg was intrigued by what he read in [**146**] and thought this might be a good topic to study. Being a mechanical engineer, he did not find the algebraic geometry of the 1980s to be helpful. Instead, Sederberg turned to the older literature on elimination, including a 1908 paper of Dixon entitled *The eliminant of three quantics in two independent variables* [**133**]. In this paper, Dixon writes:

> I now go on to the discussion of the eliminant of three quantics of degree $m+n$, in which the highest power of x and y respectively are x^n and y^m, each quantic being of the form
>
> $$\sum A_{rs}x^r y^s \quad (r=0,1,2,\ldots,n;\ s=0,1,2,\ldots,m). \quad [\mathbf{133}, \S 9]$$

The Newton polytope is the rectangle $n\Delta_1 \times m\Delta_1$, which for the polynomials (1.31) is the square



Hence we are dealing with sparse elimination. In 1984, Sederberg, Anderson and Goldman published *Implicit Representation of Parametric Curves and Surfaces*, where the introduction notes that

> ... the general theory of elimination is a lost art, with obscure references. [**312**]

The goal of their paper is to use classic ideas of Sylvester, Cayley and Dixon to implicitize certain parametric curves and surfaces.

For a curve parametrization

$$x = \frac{a(t)}{c(t)}, \quad y = \frac{b(t)}{c(t)},$$

where a, b, c are relatively prime polynomials of degree n, the authors explain how the Sylvester formula for the classical resultant gives the implicit equation

$$F(x,y) = \mathrm{Res}_t(c(t)x - a(t), c(t)y - b(t)) = 0 \tag{1.32}$$

when the parametrization is generically injective.

For surfaces, they turn to Dixon's paper, where the goal is to find the "eliminant of quantics ϕ, ψ, χ." When ϕ, ψ, χ have Newton polygon $n\Delta_1 \times m\Delta_1$, this is the sparse $\mathcal{A}$-resultant $\mathrm{Res}_{\mathcal{A}}(\phi, \psi, \chi)$ for $\mathcal{A} = (n\Delta_1 \times m\Delta_1) \cap \mathbb{Z}^2$. In [**133**], Dixon gives three methods to express this resultant as a determinant.

For a bilinear surface parametrization

$$\begin{aligned} x &= a_0 + a_1 s + a_2 t + a_3 st \\ y &= b_0 + b_1 s + b_2 t + b_3 st \\ z &= c_0 + c_1 s + c_2 t + c_3 st, \end{aligned}$$

we move x, y, z to the right hand side of each equation, and then Dixon's third method leads to the implicit equation

$$F(x,y,z) = \det \begin{bmatrix} a_0 - x & a_1 & a_2 & a_3 & 0 & 0 \\ b_0 - y & b_1 & b_2 & b_3 & 0 & 0 \\ c_0 - z & c_1 & c_2 & c_3 & 0 & 0 \\ 0 & a_0 - x & 0 & a_2 & a_1 & a_3 \\ 0 & b_0 - y & 0 & b_2 & b_1 & b_3 \\ 0 & c_0 - z & 0 & c_2 & c_1 & c_3 \end{bmatrix} = 0,$$

which is quadratic in x, y, z. This determinant appears in [**312**, 5.1]. In general, Dixon's method gives a $6mn \times 6mn$ determinant—see [**105**, Exercise 7.2.10].

The paper [**312**] also uses Dixon's first method for the resultant, which gives a $2mn \times 2mn$ determinant based on a variant of the Bézoutian from Section 1.1. The entries of the matrix are linear in x, y, z, so that the implicit equation has degree $2mn$. In the bicubic case, $n = m = 3$ gives a surface of degree 18. This number was expected to be 18 because of ray tracing, but [**312**] was the first geometric modeling paper to give a general formula for the degree of the surface. We will discuss a modern version of the degree formula in Section 3.2. All of the resultants appearing in [**312**] are given by determinants. This is rare, since most determinants giving resultants have extraneous factors, similar to Proposition 1.7. In the unmixed case, Sylvester-style representations (e.g., Dixon's third method) are classified in [**165**, Proposition 13.2.2], and [**128**] shows that each of these has a Bézout-style representation (e.g., Dixon's first method).

A Preview of the 21st Century. To give a hint of what has happened in geometric modeling since 1984, we end with a theorem from a 2016 paper of Botbol and Dickenstein [**41**]. Suppose that we have a parametrization

$$x = \frac{f_1(s,t)}{f_0(s,t)}, \quad y = \frac{f_2(s,t)}{f_0(s,t)}, \quad z = \frac{f_3(s,t)}{f_0(s,t)} \tag{1.33}$$

and want to find the implicit equation. Our only assumption on f_0, f_1, f_2, f_3 is that they are relatively prime.

One approach would be to use a sparse resultant based on the exponent vectors of the f_i. But sparse resultants are not easy to compute. So instead, we replace the resultant with a matrix of polynomials similar to what we did in Proposition 1.7 and Theorem 1.8. But those matrices were built from the coefficients of the polynomials, while here, we will use *syzygies*. Recall that a syzygy on f_0, f_1, f_2, f_3 consists of polynomials h_0, h_1, h_2, h_3 satisfying

$$h_0 f_0 + h_1 f_1 + h_2 f_2 + h_3 f_3 = 0. \tag{1.34}$$

To reduce to a finite set of syzygies, let $P \subseteq \mathbb{R}^2_{\geq 0}$ be a lattice polygon that contains the Newton polytopes of f_0, f_1, f_2, f_3. The set of all syzygies (1.34) where the Newton polytopes of the h_i are contained in $2P$ is a finite dimensional vector space that can be computed by linear algebra.

A basis of such syzygies gives polynomials

$$\mathcal{L}^{(l)} = h_0(s,t)^{(l)} w + h_1^{(l)}(s,t)x + h_2^{(l)}(s,t)y + h_3^{(l)}(s,t)z, \quad l = 1, \ldots, N,$$

where w, x, y, z are homogeneous coordinates on $\mathbb{P}^3$. Now rewrite each polynomial in the form

$$\mathcal{L}^{(l)} = \sum_{(i,j) \in 2P \cap \mathbb{Z}^2} \ell^{(l)}_{(i,j)}(w,x,y,z)\, s^i t^j$$

for linear forms $\ell^{(l)}_{(i,j)}(w,x,y,z)$. When we arrange the $\ell^{(l)}_{(i,j)}$ into a matrix M of linear forms, we have the following result from [**41**].

THEOREM 1.18. *In nice cases, the rank of the matrix M drops precisely when evaluated at the points in the closure in $\mathbb{P}^3$ of the surface parametrized by* (1.33).

There are a lot questions to ask here. What do we mean by "nice cases"? And why syzygies? All will be answered in Section 3.2. An especially intriguing feature

of Theorem 1.18 is the replacement of resultants with rank conditions on a matrix, an idea used by van der Waerden in 1926. But Theorem 1.18 is only one of many things happening in 21st century elimination theory, as you will soon see.

1.3. Elimination Theory in the 21st Century

by Carlos D'Andrea

Among the thousands of challenges that this century poses to science we can mention:

- *The average citizen should understand* (*and get excited by!*) *the latest results in science*, which explains the increasing effort towards popularization of science and scientists done by the media lately, jointly with a more active involvement of the scientific community in this task.
- *Research should be more interdisciplinary*, this trend probably being pushed by the fact that more and more funds are available in the applied-close-to-the-real-world problems, making all of us shift our gears in that direction.
- *Computations should be fast*, as we have increased enormously our computational capacities in the last decades, and still hope for improving them in the next years to come.

These challenges have definitely impacted all areas of mathematics, particularly those of a computational nature. For instance, in the last decades we witnessed the birth and growth of a large number of software packages in computational commutative algebra and algebraic geometry [**1**, **20**, **121**, **177**, **255**].

Elimination theory was not exempt from this trend, and the last few years saw us immersed in both bringing interesting (cool!) problems from chemistry, medicine, biology to our research as well as proposing new theoretical models and theorems coming from these disciplines, see for instance [**336**].

When it comes to fast computation, the "cost" (complexity) of elimination has been already studied all along the last century, going well back to the beginning of the 1900s with the work of Emmy Noether. Unfortunately, it is well known (see for instance [**198**] for a recent treatment of the topic) that this cost is at least exponential in the standard model of computation, and can get worse.

As problems coming from the real world have to be modelled and solved in a reasonable time frame, this bad result from theoretical complexity led us to consider changing the "grand picture" of chasing universal solvers for all kinds of problems in elimination theory to either *relax the input* (and expect the output to be relaxed too!), or *change the computational model* to put the relaxation there.

The 21st century is still unfolding, so as a disclaimer, we are not going to point in a general (or several) direction(s) and make prophecies on where elimination theory will be 80 years from now, because with very high probability we would fail. Instead, we are going to highlight directions that have been explored successfully already within this century, and within each category, we will choose one or two papers as sampling of how things work and progress in these areas. So, instead of having a full meal, what you will find here is a selection of some kind of Spanish "tapas" into the 21st century. We hope you enjoy them.

Relaxing the Input. We will explore several ways of doing this.

Tropical Geometry. A very easy way of getting rid of "exponential" objects is by computing their logarithms. This idea lies at the core of tropical geometry: let $V \subseteq (\mathbb{C}^*)^n$ a variety. Its "amoeba" is defined as the image of V under the logarithmic map $\mathrm{Log}_t : (\mathbb{C}^*)^n \to \mathbb{R}^n$, $\mathrm{Log}_t(z_1, \ldots, z_n) = (\log_t |z_1|, \ldots, \log_t |z_n|)$, where $t \in \mathbb{R}_{>0}$. In principle, the image of V under this map can become way more complicated to describe than V itself (for instance, if V is analytic or algebraic, its image is not necessarily an object of the same category).

However, as $t \to +\infty$, amoebas get contracted around some asymptotic directions which will define "the skeleton" of the *tropical variety* of V. The limit will be a union of pieces of linear varieties which still contain important information from the original variety, and are easier and fast to compute, which is what we are looking for. Tropical geometry takes its name from "tropical algebra", a classical construction which already existed in 20th century mathematics, which now comes into the picture as some kind of "base semi-ring" of tropical varieties. The reader will find in [**46**, **266**] and the references therein an introduction and state of the art of this fabulous and very active area of mathematics, which now has its own code **14Txx** in the Mathematics Subject Classification of the AMS.

It should not be surprising that also in tropical geometry there is a notion of "elimination", and that it is of course connected to the "tropicalization" of classical elimination theory. Our first "tapa" consists essentially of displaying such a result. Indeed, in [**338**] Bernd Sturmfels and Jenia Tevelev state the "fundamental theorem" on morphisms of tropical varieties. Their main result is the following:

THEOREM 1.19 (Theorem 1.1 in [**338**]). *Let $\alpha : \mathbb{T}^n \to \mathbb{T}^d$ be a homomorphism of tori and let $A : \mathbb{Z}^n \to \mathbb{Z}^d$ be the corresponding linear map of lattices of one-parameter subgroups. Suppose that α induces a generically finite morphism from X onto $\alpha(X)$. Then A induces a generically finite map of tropical varieties from trop(X) onto trop$(\alpha(X))$.*

Some words to explain the language and notation of this statement: the torus $\mathbb{T}^n$ here is essentially $(\mathbb{C}^*)^n$, and X is an algebraic subvariety of $\mathbb{T}^n$. It is known that (see for instance [**106**]) any homomorphism of tori as above is encoded by a linear map, in this case the map is denoted by A. The operator *trop* in the statement of Theorem 1.19 is the "tropicalization" of a variety which we have sketched when we introduced the concept of a tropical variety: replacing the variety by this union of pieces of linear varieties jointly with some less subtle structure. This theorem states that "tropicalizing" the image of a variety under a homomorphism of tori is essentially the same as tropicalizing the variety first, and then applying the linear map A. So the algebraic map α gets "simplified" to the linear map A.

Apart from being of interest on its own, this result has been used to extend and simplify the tropicalization of classical objects in elimination theory such as resultants. This work has been tackled by Anders Jensen and Josephine Yu in [**220**], which is our second "tapa" in this area of elimination and tropical geometry. We reproduce here the abstract of this paper:

> We fix the supports $\mathcal{A} = (A_1, \ldots, A_k)$ of a list of tropical polynomials and define the tropical resultant $\mathcal{TR}(\mathcal{A})$ to be the set of choices of coefficients such that the tropical polynomials have a common solution. We prove that $\mathcal{TR}(\mathcal{A})$ is the tropicalization of the algebraic variety of solvable systems and that its dimension

> can be computed in polynomial time. The tropical resultant inherits a fan structure from the secondary fan of the Cayley configuration of $\mathcal{A}$, and we present algorithms for the transversal of $\mathcal{TR}(\mathcal{A})$ in this structure. We also present a new algorithm for recovering a Newton polytope from the support of its tropical hypersurface. We use this to compute the Newton polytope of the sparse resultant polynomial in the case when $\mathcal{TR}(\mathcal{A})$ is of codimension 1. Finally we consider the more general setting of specialized tropical resultants and report on experiments with our implementations.

In this excerpt we find phrases like "tropical polynomials" which we have not defined here (and we will not), but essentially what we learn by reading the description of the main results in [**220**] is that the "tropical resultant" acts as a resultant of these tropical polynomials, and turns out to be the tropicalization of the classical resultant. This is of interest because it shows that one can compute these tropical resultants faster.

What about discriminants? These objects are also at the heart of elimination theory, with very mysterious properties which still have to be understood and explored. Their computation is still a challenge even for the simplest cases. But tropical geometry has can be of help for this task. Indeed, in the "tapa" [**130**], Eva-Maria Feichtner, Alicia Dickenstein and Bernd Sturmfels succeeded in explaining what the tropicalization of the discriminant is. Their main result is the following:

THEOREM 1.20 (Theorem 1.1 in [**130**]). *For any matrix A, the tropical A-discriminant $trop(X_A^*)$ equals the Minkowski sum of the co-Bergman fan $B^*(A)$ and the row space of A.*

To make this statement more precise (background and proofs of course can be consulted in [**130**]), any discriminant can be codified by a matrix with integer entries (in the case of Theorem 1.20 it is the matrix A). To any such integer matrix A one can associate a projective toric variety X_A, and the dual of this variety, denoted by X_A^*, is the one defining the discriminant associated to this data. The map *trop* is the same as above, the one which "tropicalizes". Finally, the *co-Bergman fan* $B^*(A)$ is just the tropicalization of the kernel of A. So, Theorem 1.20 essentially states that the tropicalization of a very complicated object (the A-discriminant) can be easily computed as the Minkowski sum of the row space of the matrix A and the tropicalization of its kernel. With this result in hand, the authors give in the same paper (Theorem 1.2) an explicit description of the initial monomials of the A-discriminant, shedding light into a problem that seemed hard to approach for decades.

"Tailored" Resultants. The main reason why computing resultants is a hard problem from a complexity point of view is because the number of monomials of a given degree is "big", in the sense that it grows exponentially as the number of variables or the degree increases. It is even worse, the number of solutions of a "generic" system of n polynomials in n unknowns with fixed degree is also exponential with respect to the degree, and resultants codify somehow these solutions. In an attempt to avoid this situation, different kinds of "resultants" appeared in the last decades, all of them essentially associated to a projection of an incidence variety as in (1.23) from Section 1.2, but replacing the projective space $\mathbb{P}^n$ with another compact variety, and hence having the Elimination Theorem be valid for this kind of variety, hopefully making the new resultant easier to compute. This approach has been very fruitful both from a theoretical point of view as well for applications,

and is still being explored these days. We will only present few "tapas" in this section, as there are several important and interesting results in a lot of directions. The reader will find the references at the beginning of each title below.

Among the results obtained here we can mention:

- Reduced resultants: Even though this operator is quite old (from the previous century), it became popular after being introduced by Oscar Zariski in [**367**] to prove some properties of multiplicities of points, and then treated more systematically by Jean-Pierre Jouanolou in [**228**, Section 5.3.29]. This is the case when all the input polynomials are generic with the property of "having already a common zero". The variety $\mathbb{P}^n$ is replaced by the blow-up of the same projective space at a point. This common zero can also be considered as "fat". Then the blow-up would be with respect to this multiple point (the defining ideal of the fat point).
- Anisotropic resultants [**225**, **227**, **228**]: These are resultants of "weighted" homogeneous polynomials, i.e., instead of declaring that the degree of the monomial x_i is equal to 1, one sets $\deg(x_i) = d_i$ for positive integers $d_1, \ldots, d_{n+1}$. The role of projective space $\mathbb{P}^n$ is now replaced by the "weighted projective space" (as defined in [**106**]) $\mathbb{P}(d_0, \ldots, d_n)$.
- Sparse resultants [**120**, **165**, **231**, **335**]: We encountered these in Section 1.2. The variety replacing $\mathbb{P}^n$ is a toric variety (or a toric cycle, see [**120**]) associated to the monomials appearing in the expansion of the input polynomials. This is why the name "sparse" arises, as the degree or the number of variables may be high, but only few monomials may appear in each of the polynomials. The complexity of elimination has been shown to be directly related to the number of monomials appearing in the input as well as the number of solutions of generic sparse systems which only use these monomials, and hence the size of these resultants is (at least theoretically) smaller than expected.
- Parametric resultants [**62**]: This resultant can be regarded as a generalization of sparse elimination, instead of considering a polynomial as a "sum of monomials" we will consider $f_i = \sum_{j=1}^{N} c_{ij} g_j(x)$, with $\{g_j(x)\}_{1 \leq j \leq N}$ being a predetermined family of polynomials. In the classical and sparse case this family would just encode the monomials appearing in the expansion of each f_i. Under suitable conditions on this family, it can be shown that a variety which is parameterized by the $g_j(x)$'s can be used to replace the classical projective space in the elimination scheme.
- Residual resultants [**52**, **63**]: This case is similar to the parametric one just described, but instead we impose suitable conditions on the family $\{g_j(x)\}_{1 \leq j \leq N}$, for instance to have a "nice" set of zeroes, a complete intersection. In this case, it is a generalization of the reduced resultants described above when there is a common (non-multiple) zero of the input system.The variety replacing the classical projective space in the elimination scheme is now replaced by the blow-up of this space along the zeroes of this family.

Of course there are other resultants which have appeared in the picture in the last years: determinantal ([**53**]), developed ([**234**]), equivariant ([**65**]),... This is a very fast moving area with plenty of room for implementation, discovery and excitement.

Implicitization Matrices. Manipulating matrices is at the core of elimination theory: matrices appear when we want to make explicit some (linear) phenomenon, and the classical connection with matrices and commutative algebra was done via determinants, the most important object in elimination theory. Determinants or minors appear most of the time to measure the failure of a given system to be as "generic" as it was supposed to be. From the very beginning of the history of elimination theory, determinants have been used to come back from the "matrix kingdom" to commutative algebra, and a lot of progress in this area was done towards the goal of finding nice determinantal formulæ for computing resultants and eliminants, to give an example of this interaction.

This approach, however, underwent an important change in the last years with the surge of numerical linear algebra. Now we can compute ranks and lots of other invariants of huge matrices even with approximated numerical input with very little effort. So these ideas should be adapted and used in computational algebra too, particularly in elimination theory. We saw a glimpse of this in the *Preview of the 21st Century* in Section 1.2.

The tapa chosen to illustrate this topic is the paper by Laurent Busé [**55**]. Let us illustrate it with the following example: consider the (affine) parametrization of the twisted cubic in affine space $\mathbb{A}^3$, $t \mapsto (t, t^2, t^3)$. If we denote by (x, y, z) the coordinates of this space, it is well known that the implicit equations of the image are given by the maximal minors of the 2×3 matrix

$$M(x, y, z) = \begin{bmatrix} 1 & x & y \\ x & y & z \end{bmatrix}. \tag{1.35}$$

This implies the following straightforward consequence: a point with coordinates $(p_1, p_2, p_3) \in \mathbb{A}^3$ lies in the twisted cubic if and only if $\operatorname{rank}\big(M(p_1, p_2, p_3)\big) \leq 1$. Matrices like (1.35) are an example of what in [**55**] are called "implicitization matrices". They should be "easy" to build (for instance, the entries should be linear), and also easy to evaluate. Moreover, they should be sensitive to numerical data, i.e., when dealing with approximate data, the ansatz should be sensitive to it. But the main property of these implicitization matrices is the fact that they drop rank after being evaluated if and only if the point of evaluation belongs to the image of the parameterized variety. The new "philosophy" here is the following: instead of considering all the maximal minors of $M(x, y, z)$ as the generators of the curve, just take the matrix itself to codify that information! It is easier to codify, and also easy to check for instance membership to the curve.

In [**55**] (see also [**56**]) there are methods to compute implicitization matrices for any rational map $\mathbb{A}^n \dashrightarrow \mathbb{A}^m$. These matrices are the result of computing a kernel of some kind of "Sylvester map", and as a result they end up being linear, as the example above. It should be noted that these implicitization matrices are not a generalization of the structure of $M(x, y, z)$ above, which essentially states something stronger, namely the fact that the twisted cubic is a determinantal variety. For example, applying the methods in [**55**] to the parametrization of the twisted cubic, one obtains matrices like the following:

$$\begin{bmatrix} x+y+z & x+y & x \\ z-1 & y-1 & x-1 \end{bmatrix},$$

or

$$\begin{bmatrix} 2x+2y & -2x-y & 2x+2y+2z & -3x-2y-z & -x & 2x \\ y-1 & 1 & z-1 & 3/2 & 1/2 & x-1 \\ 0 & y-1 & 0 & z-1 & x-1 & 0 \end{bmatrix}.$$

And here is what an implicitization matrix of the parametrization of the unit sphere in $\mathbb{A}^3$ looks like:

$$\begin{bmatrix} y & y+z-x+1 & x-y-1 & 0 \\ y & 2y-2x & x-y+z-1 & -y \\ y-x-1 & -2x & 2x & z \end{bmatrix}.$$

A key feature of these implicitization matrices is the fact that all they are actually more a representation of the parameterization itself, and not only of its image (in this case, a curve in 3-dimensional space). So, they not only give us a way to test for "membership" but also provide with a tool to compute inverse images of the parametrization at a given point in the image.

One may consider implicitization matrices as a generalization of the use of resultant matrices to solve overdetermined polynomial systems (again an inversion problem). This can be done because resultant matrices are built from syzygies, the trivial ones, the Koszul syzygies. In general, looking for implicitization matrices demands studying the syzygies of the input.

Another key feature of representation matrices is: their entries are linear in the coordinates of the target space, and testing rank should be fast when dealing with specialized coefficients. Moreover, if the ground field is $\mathbb{R}$ ("the" ground field for graphics and computer aided geometric design) one can use classical methods of numerical analysis to present these matrices in a more canonical and easier form (like the use of singular value decomposition). For all this, the use of implicitization matrices has been shown to be well suited for a lot of problems in graphics and computer aided geometric design. Operations like the study of birationality of a map, computing inverses, and intersections of varieties encoded this way can be done with the aid of these tools in a very efficient way.

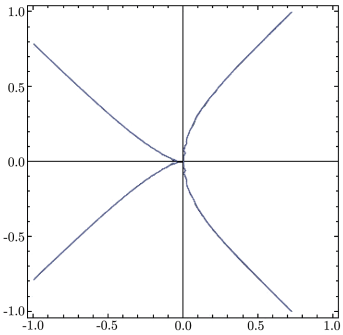
See Section 3.2 for further discussion of implicitization matrices.

Syzygies and Rees Algebras. These topics display a very nice and interesting interaction between algebraic geometers and the geometric modelers introduced in Section 1.2. In the 1990s, a series of experiments made by Tom Sederberg and his collaborators showed something interesting: one can compute the defining polynomial of a parametric rational curve as a determinant of a matrix whose entries are coefficients of what they called *moving curves* that follow the parametrization. In algebraic language, these moving curves are polynomial relations of the input polynomials defining the parametrization. The intriguing feature is that the more singular the curve is (i.e., the more singular points it has), the smaller the size of the determinant. For instance, the following result appears in [**314**]:

THEOREM 1.21. *The implicit equation of a quartic curve with no base points can be written as a* 2×2 *determinant. If the curve doesn't have a triple point, then each element of the determinant is quadratic. Otherwise, one row is linear and one row is cubic.*

To illustrate this result, we consider the following examples.

EXAMPLE 1.22. Set $a(s,t) = s^4 - t^4$, $b(s,t) = -s^2t^2$, $c(s,t) = st^3$. These polynomials define a parametrization $\Phi : \mathbb{P}^1 \to \mathbb{P}^2$, with implicit equation given by $F(x,y,z) = z^4 - y^4 - xyz^2$. Here is an affine picture of the curve over the reals when $x = 1$:



From the shape of $F(x,y,z)$, it is easy to show that $(1,0,0)$ is a point of multiplicity 3 of this curve, and it is also easy to verify that

$$\mathcal{L}_{1,1}(s,t,x,y,z) = sz + ty$$

is a "moving line" (a syzygy of (a,b,c)) which follows Φ. The reader will now easily check that the following moving curve of bidegree $(3,1)$ also follows Φ, where "following" here means that when we replace x,y,z with a,b,c, the expression below vanishes identically:

$$\mathcal{L}_{1,3}(s,t,x,y,z) = s(y^3 + xz^2) + tz^3.$$

Note that $\mathcal{L}_{1,3}(s,t,x,y,z)$ can be regarded as a syzygy on a^3, ac^2, c^3, elements of the cube of the ideal generated by the input parametrization. The 2×2 matrix claimed in Theorem 1.21 for this case is made with the coefficients of both $\mathcal{L}_{1,1}(s,t,x,y,z)$ and $\mathcal{L}_{1,3}(s,t,x,y,z)$ as polynomials in s,t:

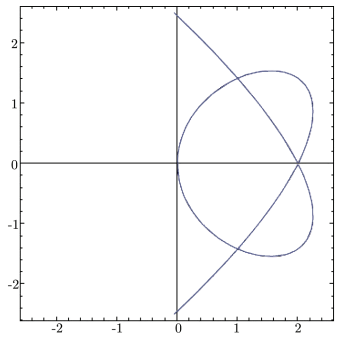
$$\begin{bmatrix} z & y \\ y^3 + xz^2 & z^3 \end{bmatrix}. \tag{1.36}$$

◁▷

EXAMPLE 1.23. We reproduce here Example 2.7 in [**96**]. Consider

$$a(s,t) = s^4,\ b(s,t) = 6s^2t^2 - 4t^4, c(s,t) = 4s^3t - 4st^3.$$

This input defines a quartic curve with three nodes, with implicit equation given by $F(x,y,z) = z^4 + 4xy^3 + 2xyz^2 - 16x^2y^2 - 6x^2z^2 + 16x^3y$.



The following two moving conics of degree 1 in s, t follow the parametrization:

$$\begin{aligned}\mathcal{L}_{1,2}(s,t,x,y,z) &= s(yz-xz)+t(-z^2-2xy+4x^2)\\ \tilde{\mathcal{L}}_{1,2}(s,ts,x,y,z) &= s(y^2+\tfrac{1}{2}z^2-2xy)+t_1(xz-yz).\end{aligned}$$

As in the previous example, the 2×2 matrix of the coefficients of these moving conics

$$\begin{bmatrix} yz-xz & -z^2-2xy+4x^2 \\ y^2+\frac{1}{2}z^2-2xy & xz-yz \end{bmatrix} \tag{1.37}$$

is the matrix claimed in Theorem 1.21. ◁▷

It is known in general that one can compute the implicit equation of a curve of degree d from the determinant of a $d\times d$ matrix containing the coefficients of syzygies ("moving lines") of the input data. This matrix can be built from the *Bézout matrix* of the parametrization (see [**61**, Proposition 6.1] for details on this construction). For $d=4$, this means that the implicit equation of a quartic can be computed from the determinant of a 4×4 matrix. Theorem 1.21 above essentially shrinks the size of the matrix in half. Is there an analogue of this result for higher values of d?

The *method of implicitization by using moving curves and surfaces* introduced by Sederberg and his collaborators in [**314**, **315**] and subsequent articles [**313**, **369**] essentially explored this question, and lots of partial results, especially using moving conics and quadrics, were obtained at the time. Among them we can mention [**60**, **77**, **94**, **100**, **108**].

The main breakthrough in this area was provided by David Cox, who showed at the beginning of this century in the tapa [**96**], that a good way to understand all these moving lines, conics, cubics, and their interactions, is by taking a look at the kernel $\mathcal{K}$ of the following morphism of rings:

$$\begin{aligned} k[s,t,x,y,z] &\longrightarrow k[s,t,e]\\ s,t &\longmapsto s,t\\ x,y,z &\longmapsto xe, ye, ze. \end{aligned} \tag{1.38}$$

Here, k is the ground field where the parametrization is defined, and e a new variable. Indeed, it is shown in [**96**, Nice Fact 2.4] (see also [**64**] for the case when Φ is not generically injective) that all the moving curves that follow the parametrization Φ are bihomogeneous elements of $\mathcal{K}$, and they generate this ideal.

Let us say some words about the map (1.38). Denote with $I\subseteq k[s,t]$ the ideal generated by $a(s,t)$, $b(s,t)$, $c(s,t)$. The image of (1.38) is actually isomorphic to $k[s,t][e\,I]$, which is the *Rees algebra* of I. By the Isomorphism Theorem, we then get that $k[s,t,x,y,z]/\mathcal{K}$ is isomorphic to the Rees algebra of I. This is why the generators of $\mathcal{K}$ are called the *defining equations* of the Rees algebra of I. The Rees algebra that appears in the method of moving curves corresponds to the blow-up of $\mathbf{V}(I)$, the variety defined by I. Geometrically, it is just the blow-up of the empty space (the effect of this blow-up introduces only torsion), but this geometrical perspective should explain somehow why moving curves are sensitive to the presence of complicated singularities. As of today it is not well understood why the description of $\mathcal{K}$ actually gets much simpler if the singularities of the curve are more entangled.

Indeed, it has been shown in [**54**], by unravelling some duality theory, that for any proper parametrization of a curve of degree d having a moving line of

degree 2 and only cusps as singular points, the kernel $\mathcal{K}$ has $\frac{(d+1)(d-4)}{2}+5$ minimal generators. On the other hand, in joint work with Teresa Cortadellas [**90**] (see also [**247**]), we have shown that if there is a moving line of degree 2 and there is a point of higher multiplicity, then the number of generators drops to $\lfloor\frac{d+6}{2}\rfloor$, i.e., the description of $\mathcal{K}$ is simpler in this case. In both cases, these generators can be made explicit, see [**54**, **90**, **247**].

Further evidence supporting this claim is what we already know for the case when there is a moving line of degree 1, which was one of the first cases worked out by several authors [**54**, **89**, **101**, **210**]. It turns out (cf. [**89**, Corollary 2.2]) that $\mu = 1$ if and only if the parametrization is proper (i.e., generically injective) and there is a point on the curve with multiplicity $d-1$, which is the maximal multiplicity a point can have on a curve of degree d. In this case, $\mathcal{K}$ has $d+1$ minimal generators, of bidegrees

$$\{(0,d),\ (1,1),\ (1,d-1),\ (2,d-2),\ldots,(d-1,1)\},$$

and one can show easily that the determinant of the 2×2 matrix formed by collecting the coefficients of the forms of bidegrees $(1,1)$ and $(1,d-1)$ in this family as it was done in (1.36) and (1.37), gives the implicit equation of the curve. So, for the case $\mu = 1$ we can compute implicit equations with determinants of 2×2 matrices, independent of the value of d.

In both cases ($\mu = 1$ and $\mu = 2$), explicit elements of a set of minimal generators of $\mathcal{K}$ can be given in terms of the input parametrization. But in general, very little is known about how many they are, what their bidegrees are, and how to compute them. This is a very active and promising are of research, with several results already obtained [**91**, **92**, **96**, **101**, **247**, **260**, **267**] and yet more to be discovered in the following years...

This topic will be further elaborated in Sections 3.2 and 3.3.

Changing the Computational Model. The rigidity of Turing's model of computation (which is still the standard to this day) combined with the fact that, to study the complexity of an algorithm in this model, we must consider all the possible scenarios and most of the time take into account the worst case of each of them, sometimes leads to this unpleasant exponential number of steps that haunts elimination theory. To overcome this problem, alternative models and algorithms have been proposed since the very beginning of the computer era. Several of them are still being explored and implemented, as they turn out to be very useful for practical computations. We will go through some of them, highlighting a few pearls in elimination theory which have appeared in this century.

Probabilistic algorithms. These algorithms make at some point "random" choices which have (most of the time) very high probability of success, and continue carrying on with their task. In this way, one can avoid a large number of operations and, with some luck, get to the right answer. An example of such an algorithm with strong implications in elimination theory was given by Gabriela Jeronimo, Teresa Krick, Juan Sabia, and Martín Sombra in [**221**]. Indeed, in the abstract of this article we read:

> We present a bounded probability algorithm for the computation of the Chow forms of the equidimensional components of an algebraic variety. In particular, this gives an alternative procedure for the effective equidimensional decomposition of the variety, since each equidimensional component is characterized by

> its Chow form. The expected complexity of the algorithm is polynomial in the size and the geometric degree of the input equation system defining the variety. Hence it improves (or meets in some special cases) the complexity of all previous algorithms for computing Chow forms...
> ... As an application, we obtain an algorithm to compute a subclass of sparse resultants, whose complexity is polynomial in the dimension and the volume of the input set of exponents. As another application, we derive an algorithm for the computation of the (unique) solution of a generic overdetermined polynomial equation system.

Before describing roughly how this algorithm works, we will just say that the "Chow form" of an algebraic variety is a polynomial in a lot of variables which generalizes all the resultants described above.

What follows a rough description of the algorithm. Note that no "general formula" or "determinantal" or "matrix form" appears in the different steps of this procedure, as in probabilistic algorithms one should use different shoes to walk through the resolution of a problem.

Computing the Chow Form of a variety [**221**]:
Given (equations of) a d-dimensional variety $V \subseteq \mathbb{P}^n$,

(1) Intersect V with d "general" (probabilistically chosen) hyperplanes to obtain finitely many points

$$V \cap \{H_1 = \dots H_d = 0\} = \{\xi_1, \dots, \xi_D\} \subseteq \mathbb{P}^n.$$

(2) From (a "geometric resolution" of) the set $\{\xi_1, \dots, \xi_D\}$, lift via Newton's method to the Chow Form of

$$V \cap \{H_1 = \dots = H_{d-1}\}.$$

Note that the latter set should be now one-dimensional.

(3) Iterate this process $d-1$ times to finally get the Chow Form of V.

As explained in that paper, this algorithm is a probabilistic *Blum-Shub-Smale machine* over $\mathbb{Q}$ which manipulates *straight line programs*. These machines are generalizations of classical Turing's machines to the rationals in the sense that instead of manipulating only 0's and 1's, any number over $\mathbb{Q}$ is treated in the same way. As explained in the book [**33**], a BSS machine is a Random Access Machine with registers that can store arbitrary numbers and can compute rational functions at unit cost. So, for this model, representing *any* number has a very cheap cost. There are several challenges in elimination theory within the formulation of BSS machines, as we will see soon.

What are straight-line programs? These are very old friends of ours, although hard to imagine that they can have such a relevant role in modern computer algebra: a straight-line program is an algorithm (a "program") with no branches, no loops, no conditionals, no statements, no comparisons, just a sequence of basic operations. A basic straight-line program is the evaluation of a polynomial with integer coefficients at a given point: one does not have to stop at any intermediate step to compare numbers or decide to change the formula if some prescribed value is reached or avoided. Of course the hard part in designing *efficient* straight-line programs is to get fast and useful procedures which solve a given task without bifurcations. For instance, evaluating a polynomial at a given point the way we learnt in high school probably is not a good straight-line program, as using the Horner

scheme of evaluation is another of such straight-line programs which is better from a complexity point of view.

Straight-line programs have a rich and interesting history in computational algebra and algorithms. The interested reader can learn more about them in [**170**, **171**, **243**].

Another highlight in this model of computation and using similar tools was given recently by Gabriela Jeronimo and Juan Sabia in [**222**], where they succeeded in producing a straight-line program which evaluates resultants in polynomial time. The abstract of their paper states:

> We prove that the sparse resultant, redefined by D'Andrea and Sombra and by Esterov as a power of the classical sparse resultant, can be evaluated in a number of steps which is polynomial in its degree, its number of variables and the size of the exponents of the monomials in the Laurent polynomials involved in its definition. Moreover, we design a probabilistic algorithm of this order of complexity to compute a straight-line program that evaluates it within this number of steps.

To accomplish this task, the authors use very cleverly a geometric resolution of a toric variety jointly with the Newton Hensel method and Padé approximations.

Parallelization. If the resolution of a system of polynomial equations becomes exponential because the number of roots to be computed is already exponential, there is some hope that by relaxing the problem and requiring just to compute "one root", the number of steps may reduce. However, to choose one root from an algebraic system we must leave the algebraic representation that we are used to in algebra, and deal with the language and the tools that produce roots in the computational and numerical world. One of these powerful tools is Newton's method, which lies at the heart of several probabilistic algorithms like the two we just mentioned above. Both in theory and in practice, Newton's method is very fast and efficient *provided that* the Newton operator applied to the initial point converges, which seems to happen most of the time, but unfortunately it is not very easy to certify when Newton's method will succeed at a given point.

Inspired by this, we will say that a point $z_0 \in \mathbb{C}$ is an *approximate root* of $f(x) \in \mathbb{C}[x]$ if Newton's method applied to z_0 converges to a root of $f(x)$ in $\mathbb{C}$. The analogue in n dimensions is straightforward: one needs a point in $\mathbb{C}^n$, and a square system (n polynomials in n unknowns) of polynomials so that the Newton operator is also defined, and approximate roots of this system will be those n-tuples in $\mathbb{C}^n$ converging to an exact zero of the system after successive applications of the multivariate Newton operator.

In 1998 Steve Smale ([**323**]) listed 18 problems that, in his opinion, would be challenges for the 21st century. Among them, problem number 17 asks for the existence of a deterministic algorithm for computing an approximate solution of such a square system of complex polynomials in time polynomial, on the average, in the size of the input system. Like most of the problems in that list, this one received a lot of attention and important progress was done towards its resolution in the first years of the century, see [**23**, **24**, **51**]. But it was Pierre Lairez ([**249**]) who finally tackled the problem and solved it in 2017. The abstract of his paper states:

> We describe a deterministic algorithm that computes an approximate root of n complex polynomial equations in n unknowns in average polynomial time with

respect to the size of the input, in the Blum-Shub-Smale model with square root. It rests upon a derandomization of an algorithm of Beltràn and Pardo and gives a deterministic affirmative answer to Smale's 17th problem. The main idea is to make use of the randomness contained in the input itself.

Homotopy Methods. All the papers that dealt with Smale's 17th problem mentioned above used this classical mathematical idea which turns out to be very fruitful from a computational point of view: if you want to solve a very complicated system $\mathbf{f} = 0$, you can start with an "easy" system $\mathbf{g} = 0$, and then "connect" both systems with an homotopy of the type $t\mathbf{f} + (1-t)\mathbf{g} = 0$, with $t \in \mathbb{C}$ a new parameter, and then as t goes from 0 to 1 through some path over the complex numbers, then your "easy roots" will become the roots of the complicated system $\mathbf{f} = 0$. Of course a lot of complications must be overcome here. For instance, what is a "right path" to take in the homotopy? Can we reach all the roots of the system $\mathbf{f} = 0$ this way? Can we lose roots along the path? How can we do all these tasks with a computer and get the right answer?

A lot of research has been done around these problems, with very important results achieved in the last decade, and a lot of challenges remain in front of us. These homotopy methods are at the basis of a new area called *numerical algebraic geometry* [**194**], which will be explained and illustrated in detail in Chapter 2 of this book, so we are not going to add more on this topic. We only mention that with this naive idea of performing homotopies, several algebraic invariants of varieties can be computed with precision, like dimension, degree, number of irreducible components, and more.

We close this selection of "tapas of the 21st century" with a very nice result obtained by Sandra Di Rocco, David Eklund, Chris Peterson, and Andrew Sommese in [**125**], where homotopy methods were used to compute the Chern numbers of a smooth algebraic variety. This should definitely convince you that numerical algebraic geometry has already established its place in this century, and lots of interesting and useful results are appearing and will continue to appear in this area.

Conclusion. We have explored above a selection of tapas/results that should make you more like "feeling the flavor" (and not the highlights) of the state of the art in elimination theory in this century. Due to lack of space (and also personal knowledge), a lot of topics have not been covered here like differential/difference elimination theory, real elimination, direct applications of this area, or implementations. We hope that with this sampling of results, you will be convinced that there are several interesting and exciting challenges in elimination theory for the following years to come. Salud!

CHAPTER 2

Numerical Algebraic Geometry

In applications, we rarely have the luxury of knowing the exact solutions of a system of polynomial equations; approximations by floating point numbers are usually the best we can do. In fact, the coefficients of the polynomials are often floating point numbers themselves. Thus numerical issues become important when solving polynomial systems in the real world.

The first section gives simple examples to illustrate some of the numerical complications that arise and explores the power of *numerical linear algebra*. In the second section we discuss the important technique of *homotopy continuation* and give some interesting examples. In the third section, Jonathan Hauenstein will discuss some recent applications of *witness point sets*.

2.1. Numerical Issues and Linear Algebra

For someone trained in classical algebraic geometry, working with floating point numbers can be an unsettling experience. We start with the binomial theorem.

EXAMPLE 2.1. The binomial theorem gives the identities

$$\begin{aligned}(x-1)^6 &= x^6 - 6x^5 + 15x^4 - 20x^3 + 15x^2 - 6x + 1\\ &= (((((x-6)x+15)x-20)x+15)x-6)x+1.\end{aligned}$$

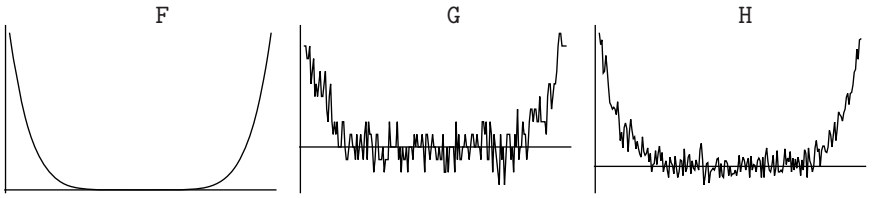
In algebra, evaluation is a ring homomorphism, so that evaluating these expressions at a real number gives the same answer. However, van Loan [**354**, p. 44] notes that

Algorithms that are equivalent mathematically may behave differently numerically.

To see what this means, we convert the above identities into three MATHEMATICA functions, where `N[...,16]` means working with 16 digit floating point numbers:

```
F[x_] = N[(x - 1)^6, 16]
G[x_] = N[x^6 - 6 x^5 + 15 x^4 - 20 x^3 + 15 x^2 - 6 x + 1, 16]
H[x_] = N[(((((x - 6) x + 15) x - 20) x + 15) x - 6) x + 1, 16]
```

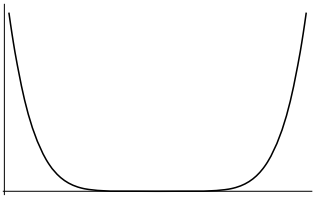
If we evaluate `F`, `G`, `H` at 200 equally spaced inputs in the interval $[.995, 1.005]$ using 16 digits of precision, we get the pictures:



Three algebraically equal expressions evaluate to very different graphs! The graph for `G` is the worst since it involves the most multiplications. But now redefine `G` to use 26 digits of precision:

```
G[x_] = N[x^6 - 6 x^5 + 15 x^4 - 20 x^3 + 15 x^2 - 6 x + 1, 26]
```

When we plot values at 200 equally spaced inputs in the interval $[.995, 1.005]$, we get the graph:



To the human eye, this is indistinguishable from the graph of $(x-1)^6$. ◁▷

In numerical algebraic geometry, it is often necessary to increase the precision automatically during certain parts of a computation. This is the idea of *adaptive precision* [**19**] that we will encounter in Section 2.2.

Our second example is taken from *Ideals, Varieties and Algorithms* [**104**].

EXAMPLE 2.2. To find the minimum and maximum values of the function $x^3 + 2xyz - z^2$ on the unit sphere $x^2 + y^2 + z^2 = 1$ in $\mathbb{R}^3$, we use Lagrange multipliers to get a system of equations in λ, x, y, z:

$$\begin{aligned} 3x^2 + 2yz &= \lambda \cdot 2x \\ 2xz &= \lambda \cdot 2y \\ 2xy - 2z &= \lambda \cdot 2z \\ x^2 + y^2 + z^2 &= 1. \end{aligned} \tag{2.1}$$

A lex Gröbner basis with $\lambda > x > y > z$ is given by the eight polynomials

$$\begin{aligned} g_1 &= \lambda - \frac{3}{2}x - \frac{3}{2}yz - \frac{167616}{3835}z^6 + \frac{36717}{590}z^4 - \frac{134419}{7670}z^2 \\ g_2 &= x^2 + y^2 + z^2 - 1 \\ g_3 &= xy - \frac{19584}{3835}z^5 + \frac{1999}{295}z^3 - \frac{6403}{3835}z \\ g_4 &= xz + yz^2 - \frac{1152}{3835}z^5 - \frac{108}{295}z^3 + \frac{2556}{3835}z \\ g_5 &= y^3 + yz^2 - y - \frac{9216}{3835}z^5 + \frac{906}{295}z^3 - \frac{2562}{3835}z = 0 \\ g_6 &= y^2z - \frac{6912}{3835}z^5 + \frac{827}{295}z^3 - \frac{3839}{3835}z \\ g_7 &= yz^3 - yz - \frac{576}{59}z^6 + \frac{1605}{118}z^4 - \frac{453}{118}z^2 \\ g_8 &= z^7 - \frac{1763}{1152}z^5 + \frac{655}{1152}z^3 - \frac{11}{288} = z\left(z^2-1\right)\left(z^2 - \frac{4}{9}\right)\left(z^2 - \frac{11}{128}\right). \end{aligned}$$

The final polynomial g_8 has roots $z = 0, \pm 1, \pm 2/3, \pm\sqrt{11}/(8\sqrt{2})$, all real and exact. As in [**104**, Example 3 of Chapter 2, §8], one can work backwards to find the 10 solutions of (2.1), again all real and exact. Evaluating $x^3 + 2xyz - z^2$ at these solutions makes it easy to determine the desired extreme values.

This example is lovely but does not represent what happens in most applications. To get a sense of what's involved, let's return to (2.1) and change the middle coefficient of $x^3+2xyz-z^2$ from 2 to nearby values such as 2.1 or 2.01. So consider $x^3+cxyz-z^2$, where c is a parameter. Finding the extrema of this function on $x^2+y^2+z^2=1$ leads to the system of equations

$$\begin{aligned} 3x^2+cyz &= \lambda\cdot 2x \\ cxz &= \lambda\cdot 2y \\ cxy-2z &= \lambda\cdot 2z \\ x^2+y^2+z^2 &= 1. \end{aligned} \tag{2.2}$$

For $c=2.1=21/10$ and $c=2.01=201/100$, a lex Gröbner basis with $\lambda>x>y>z$ still consists of eight polynomials, where $g_8=g_8(z)$ eliminates all variables but z. A surprise is that g_8 has degree 9 for these new values of c, while g_8 had degree 7 for $c=2$. The following table shows the roots of g_8 for $c=2, 2.1, 2.01$ to four decimal places:

c	first seven roots of g_8				two new roots
2	0	±1	$\pm2/3$ $=\pm.6667$	$\pm\sqrt{11}/(8\sqrt{2})$ $=\pm.2932$	
2.1	0	±1	$\pm.6592$	$\pm.3139$	$\pm14.9447\sqrt{-1}$
2.01	0	±1	$\pm.6659$	$\pm.2957$	$\pm142.2176\sqrt{-1}$

As $c\to 2$, the first seven roots approach the exact values found when $c=2$, while the two new roots approach ∞ (2.01 is ten times closer to 2 than 2.1, and the new roots are correspondingly ten times larger).

For the whole system (2.2), there are 12 solutions when $c=2.1$ or 2.01. Two of these go to ∞ as $c\to 2$, leaving 10 finite solutions when $c=2$. This analysis reveals that the original system (2.1) has solutions at ∞. ◁▷

Numerical Algebraic Geometry. Examples 2.1 and 2.2 illustrate the challenges that arise when we combine polynomials and floating point numbers. The difficulty of doing algorithmic computations with real and complex numbers has been around for a long time. Five hundred years ago, Leonardo da Vinci observed that

> Arithmetic is a computational science in its calculation, with true and perfect units, but it is of no avail in dealing with continuous quantity. [**217**, p. 201]

Numerical analysis was invented to address this difficulty, and applied to systems of polynomial equations, the result is numerical algebraic geometry.

When dealing with polynomials over $\mathbb{R}$ or $\mathbb{C}$, there are several paradigms for how to proceed. Here are brief comments about two, as applied to the system (2.2):

- In [**329**], Stetter uses the language of *empirical coefficient*, where "$c=2$ within ±0.1" means that $|c-2|\le 0.1$ is likely, $0.1\le|c-2|\le 0.2$ is possible, and $|c-2|\ge 1$ is unlikely.
- In [**324**], Sommese and Wampler regard c as a *parameter* with a generic behavior (12 solutions). Then *homotopy continuation* is used for a specific value of the parameter such as $c=2$ (10 solutions).

We will explore the Sommese and Wampler approach in Section 2.2 when we discuss homotopy continuation.

We should also mention the powerful framework created by Blum, Shub and Smale. Their methods involve α-theory and what are now called BSS machines. See [**33**] for the full story and [**196**, **249**] for applications to algebraic geometry.

Linear Algebra and Algebraic Geometry. It is time to do some algebraic geometry in a numerical context. Here, we will focus on the interaction between linear algebra and algebraic geometry. In Section 2.2, we will turn our attention to homotopy continuation.

The basic idea is that when a polynomial system has finitely many solutions over $\mathbb{C}$, basic tools from linear algebra (bases, eigenvalues, eigenvectors) tell us a lot about the solutions. We first review the basic theory and then discuss what happens numerically. The starting point is the following classic result.

THEOREM 2.3 (Finiteness Theorem). *Let $I \subseteq \mathbb{C}[x_1, \dots, x_n]$ be an ideal. Then the following are equivalent:*

(1) $\mathbf{V}(I) \subseteq \mathbb{C}^n$ *is finite.*
(2) $\dim \mathbb{C}[x_1, \dots, x_n]/I < \infty$.
(3) *I has a Gröbner basis G such that for all i, there is $g \in G$ with $\operatorname{LT}(g) = x_i^{m_i}$.*
(4) *Every reduced Gröbner basis of I has the above property.*

Furthermore, when this happens, the cosets of the monomials

$$B = \{x^\alpha = x_1^{a_1} \cdots x_n^{a_n} \mid x^\alpha \notin \langle \operatorname{LT}(I) \rangle = \langle \operatorname{LT}(G) \rangle\}$$

form a basis of $\mathbb{C}[x_1, \dots, x_n]/I$ (necessarily finite).

An ideal I as above is *zero-dimensional* since $\mathbf{V}(I)$ consists of finitely many points and hence has dimension zero as a variety. The monomials $x^\alpha \in B$ are called *remainder monomials*. The set B has two properties:

- B is an *order ideal*: if $x^\alpha \in B$, then any monomial dividing x^α is also in B.
- B is *connected to* 1: if $x^\alpha \in B$, then either $x^\alpha = 1$ or $x^\alpha = x_i x^\beta$ for some x_i and some $x^\beta \in B$.

The upshot is that once we have a zero-dimensional ideal $I \subseteq \mathbb{C}[x_1, \dots, x_n]$, we get the finite-dimensional vector space $\mathbb{C}[x_1, \dots, x_n]/I$ and a nice monomial basis B.

We also have some interesting linear maps defined on this vector space.

Eigenvalues of Multiplication Maps. Since $\mathbb{C}[x_1, \dots, x_n]/I$ is a $\mathbb{C}$-algebra, a polynomial $f \in \mathbb{C}[x_1, \dots, x_n]$ gives the linear map

$$m_f : \mathbb{C}[x_1, \dots, x_n]/I \longrightarrow \mathbb{C}[x_1, \dots, x_n]/I$$

defined by $m_f([g]) = [fg]$. In 1981, Lazard [**253**] proved the following result.

THEOREM 2.4 (Eigenvalue Theorem). *Assume $\dim \mathbb{C}[x_1, \dots, x_n]/I < \infty$ and $f \in \mathbb{C}[x_1, \dots, x_n]$. Then*

$$\big\{\text{eigenvalues of } m_f\big\} = \big\{f(p) \mid p \in \mathbf{V}(I)\big\}.$$

When we set $f = x_i$, it follows that the eigenvalues of the linear map m_{x_i} are the ith coordinates of the solutions. So we have reduced solving equations to an eigenvalue problem! Here is a simple example.

EXAMPLE 2.5. The polynomial system

$$f_1 = \underline{x^2} + 2y^2 - 2y = 0$$
$$f_2 = \underline{xy^2} - xy = 0$$
$$f_3 = \underline{y^3} - 2y^2 + y = 0$$

has Gröbner basis $\{f_1, f_2, f_3\}$ with underlined leading terms for lex order with $x > y$. Now apply the Finiteness Theorem:

- There are finitely many solutions since $\text{LT}(f_1) = x^2, \text{LT}(f_3) = y^3$.
- The remainder monomials form the order ideal $B = \{1, y, y^2, x, xy\}$.
- $\mathbb{C}[x, y]/\langle f_1, f_2, f_3\rangle$ has dimension 5.

The matrices of m_x and m_y relative to $B = \{1, y, y^2, x, xy\}$ are

$$M_x = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & -2 & -2 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{bmatrix}, \quad M_y = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}.$$

The first two columns of M_x follow since m_x maps the basis elements $1, y$ to the basis elements x, xy. For the third column, $f_2 = xy^2 - xy$ implies

$$m_x([y^2]) = [xy^2] = [xy].$$

The other entries are also easy to find. Let's compute eigenvalues:

- $\text{CharPoly}_{M_x}(u) = -u^5$ implies that 0 is the only eigenvalue of m_x.
- $\text{CharPoly}_{M_y}(u) = -u^2(u-1)^3$ implies that $0, 1$ are the eigenvalues of m_y.

By the Eigenvalue Theorem, all solutions have $x = 0$ and $y = 0, 1$. Hence the solutions are $(0, 0), (0, 1)$, found by linear algebra! ◁▷

The Eigenvalue Theorem can be modified to take multiplicities into account. Suppose $I = \langle f_1, \ldots, f_s\rangle$, so that $p \in \mathbf{V}(I)$ has a multiplicity $m(p, I)$ as a solution of the polynomial system $f_1 = \cdots = f_s = 0$ (see [**105**, Definition 4.2.1]). Then Theorem 2.4 can be stated as

$$\text{CharPoly}_{M_f}(u) = \prod_{p \in \mathbf{V}(I)} (f(p) - u)^{m(p,I)} \tag{2.3}$$

(see [**105**, Proposition 4.2.9]). When f takes distinct values at the solutions, (2.3) implies that $m(p, I)$ is the multiplicity of $f(p)$ as a root of CharPoly_{M_f}. We can see this in Example 2.5, where the solutions have distinct y-coordinates. Hence $\text{CharPoly}_{M_y}(u) = -u^2(u-1)^3$ tells us that $(0, 0)$ has multiplicity 2 and $(0, 1)$ has multiplicity 3 as solutions of $f_1 = f_2 = f_3 = 0$.

Eigenvectors of Multiplication Maps. Theorem 2.4 gives the coordinates of the solutions but doesn't explain how to match them. One way is to observe that $m_{x_i} m_{x_j} = m_{x_j} m_{x_i}$ for all i, j since $\mathbb{C}[x_1, \ldots, x_n]/I$ is commutative. We know that simultaneous eigenvectors exist in this situation, and for a simultaneous eigenvector v, the corresponding eigenvalues $m_{x_i}(v) = p_i v$ give the solution $p = (p_1, \ldots, p_n)$. In general,

$$\{m_f \mid f \in \mathbb{C}[x_1, \ldots, x_n]\}$$

is a commuting family of linear maps on $\mathbb{C}[x_1, \ldots, x_n]/I$. This leads to a rich algebraic structure explored in [**241**].

Another approach is to use the dual map

$$m_f^* : (\mathbb{C}[x_1, \ldots, x_n]/I)^* \longrightarrow (\mathbb{C}[x_1, \ldots, x_n]/I)^*.$$

If M_f is the matrix of m_f with respect to the basis B, then

$$\text{eigenvectors of } m_f^* \longleftrightarrow \text{eigenvectors of } M_f^T \longleftrightarrow \text{left eigenvectors of } M_f$$

since the transpose M_f^T is the matrix of m_f^* with respect to the dual basis of B. In 1988, Auzinger and Stetter [**12**] showed that a solution $p \in \mathbf{V}(I)$ gives a left eigenvector of M_f as follows.

THEOREM 2.6 (Eigenvector Theorem). *Assume* $\dim \mathbb{C}[x_1, \ldots, x_n]/I < \infty$ *with monomial basis* $B = \{x^{\alpha_1}, \ldots, x^{\alpha_N}\}$, *and for* $f \in \mathbb{C}[x_1, \ldots, x_n]$, *let* M_f *be the matrix of* m_f *with respect to* B. *Then:*

(1) *Every* $p \in \mathbf{V}(I)$ *satisfies*

$$(p^{\alpha_1}, \ldots, p^{\alpha_N})M_f = f(p)(p^{\alpha_1}, \ldots, p^{\alpha_N}), \text{ where } p^{\alpha_i} = x^{\alpha_i}(p).$$

(2) *If* f *takes distinct values on* $\mathbf{V}(I)$ *and* M_f *is non-derogatory, then* (1) *gives* all *eigenvectors of* M_f *up to scalar multiplication.*

Recall that a matrix M_f is *non-derogatory* if its minimal polynomial equals its characteristic polynomial, or equivalently, each eigenvalue has a unique Jordan block in the Jordan canonical form of M_f.

EXAMPLE 2.7. Consider the polynomial system from Example 2.5, where $B = \{1, y, y^2, x, xy\}$ gives a basis of $\mathbb{C}[x, y]/\langle f_1, f_2, f_3\rangle$. One can check that M_x and M_y are derogatory, but M_{2x+3y} is non-derogatory and has eigenvalues $0, 3$ with corresponding left eigenvectors

$$\begin{array}{c} 1, \quad y, \; y^2, \quad x, \; xy \\ \quad\;\; \downarrow \qquad\; \downarrow \\ \left\{\begin{array}{ll} 0\text{:} & (1, \quad 0, \quad 0, \quad 0, \quad 0) \\ 3\text{:} & (1, \quad 1, \quad 1, \quad 0, \quad 0) \end{array}\right\} \end{array}, \text{ so the solutions are } \begin{array}{c} x, y \\ \downarrow\; \downarrow \\ \left\{\begin{array}{l} (0,0) \\ (0,1) \end{array}\right\} \end{array} \text{ by Theorem 2.6.}$$

Sometimes M_f is always derogatory—see [**97, 241**] for the algebra and geometry that govern whether or not non-derogatory multiplication maps exist. ◁▷

Numerical Linear Algebra. It is now clear that linear algebra can help solve zero-dimensional polynomial systems. But Examples 2.5 and 2.7 have very simple eigenvalues. In applications, the eigenvalues tend to be more complicated, which is where *numerical linear algebra* enters the picture.

For a vector $u \in \mathbb{C}^n$, we use the standard 2-norm $\|u\|_2 = \sqrt{\sum_{i=1}^n |u_i|^2}$. Then the 2-norm of an $n \times n$ complex matrix A is

$$\|A\|_2 = \max_{\|u\|_2=1} \|Au\|_2,$$

which records the maximum "stretch" of A. When A is invertible, we define the 2-*condition number* of A to be

$$\text{cond}_2(A) = \|A^{-1}\|_2 \cdot \|A\|_2.$$

It is easy to see that $\text{cond}_2(A) \geq 1$, with equality when A is unitary. The intuition is that when $\text{cond}_2(A) \approx 10^a$, one loses roughly a digits of accuracy when solving the linear system $A\vec{x} = \vec{b}$. Thus computing accurate solutions gets more expensive as $\text{cond}_2(A)$ increases.

EXAMPLE 2.8. The 5×5 Hilbert and Pascal matrices are

$$H_5 = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} & \frac{1}{8} \\ \frac{1}{5} & \frac{1}{6} & \frac{1}{7} & \frac{1}{8} & \frac{1}{9} \end{bmatrix}, \quad P_5 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 6 & 10 & 15 \\ 1 & 4 & 10 & 20 & 35 \\ 1 & 5 & 15 & 35 & 70 \end{bmatrix}.$$

These matrices have determinants $\det(H_5) = 3.7493 \times 10^{-12}$ and $\det(P_5) = 1$. Since Cramer's rule has the determinant in the denominator, you might expect $H_5 \vec{x} = \vec{b}$ to be harder to solve numerically than $P_5 \vec{x} = \vec{b}$. This is true, but

$$\text{cond}_2(H_5) \approx 476607.6, \quad \text{cond}_2(P_5) \approx 8517.5$$

implies that we lose 5 decimal digits of accuracy using H_5 and 3 using P_5. ◁▷

You can read more about condition numbers, including their relation to the *singular value decomposition* (SVD), in [**175**, Section 2.6].

Generic Multiplication Matrices. Since the multiplication matrices M_{x_i} can be used to solve zero-dimensional polynomial systems, it makes sense to ask how hard they are to compute accurately. We will explore a particular case based on the 2018 paper *A Stabilized Normal Form Algorithm for Generic Systems of Polynomial Equations* by Telen and Van Barel [**343**].

When $f, g \in \mathbb{C}[x, y]$ are generic of degree d, some very nice things happen:

- $B = \{x^a y^b \mid 0 \le a, b \le d-1\}$ is a basis of $\mathbb{C}[x, y]/\langle f, g \rangle$. Note that B is an order ideal with $|B| = d^2$.
- The system $f = g = 0$ has d^2 solutions in $\mathbb{C}^2$ by Bézout's Theorem.
- There are formulas for the multiplication matrices M_x, M_y.

In Example 2.5, we used a Gröbner basis to find the multiplication matrices. But here, no Gröbner basis is needed—we can compute M_x and M_y directly from f, g. The full story is explained in [**105**, Section 3.6], so we will instead just give the basic idea of how this is done.

The first step is to create a Macaulay matrix M similar to the ones encountered in Chapter 1. We express the products $x^i y^j f, x^i y^j g$ for $i + j \le d - 1$ in terms of monomials of degree $2d - 1$, where columns corresponding to monomials of B are on the right. Here is an example of M with $d = 2$ taken from [**343**].

EXAMPLE 2.9. Consider the system:

$$f = x^2 + y^2 - 2, \quad g = 3x^2 - y^2 - 2.$$

Here, $B = \{1, x, y, xy\}$, and when we express f, xf, yf, g, xg, yg in terms of the monomials of degree 3, with $1, x, y, xy$, we get the matrix (omitting zero entries)

$$M = [M_1 \mid M_2] = \begin{array}{c} \\ f \\ xf \\ yf \\ g \\ xg \\ yg \end{array} \begin{array}{c} \begin{array}{ccccccccccc} x^3 & & x^2y & & xy^2 & & y^3 & & x^2 & y^2 \end{array} \\ \left[\begin{array}{ccccccccccc} & & & & & & & & 1 & 1 \\ 1 & & & & 1 & & & & & \\ & & 1 & & & & 1 & & & \\ & & & & & & & & 3 & -1 \\ 3 & & & & -1 & & & & & \\ & & 3 & & & & -1 & & & \end{array}\right. \end{array} \begin{array}{c} \begin{array}{cccc} 1 & x & y & xy \end{array} \\ \left.\begin{array}{cccc} -2 & & & \\ & -2 & & \\ & & -2 & \\ -2 & & & \\ & -2 & & \\ & & -2 & \end{array}\right] \end{array}.$$

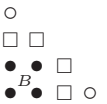
Since M_1 is an invertible 6×6 matrix, we can multiply each side by M_1^{-1} to obtain

$$(I_6 \mid M_1^{-1}M_2) = \begin{array}{c} \\ x^3-x \\ \underline{x^2y}-y \\ \underline{xy^2}-x \\ y^3-y \\ \underline{x^2}-1 \\ \underline{y^2}-1 \end{array} \begin{array}{c} \begin{array}{cccccc|cccc} x^3 & x^2y & xy^2 & y^3 & x^2 & y^2 & 1 & x & y & xy \end{array} \\ \left[\begin{array}{cccccc|cccc} 1 & & & & & & & -1 & & \\ & 1 & & & & & & & -1 & \\ & & 1 & & & & & -1 & & \\ & & & 1 & & & & & -1 & \\ & & & & 1 & & -1 & & & \\ & & & & & 1 & -1 & & & \end{array}\right] \end{array},$$

where each row label is the polynomial determined by the coefficients of the row. The underlined monomials will be explained shortly.

Multiplying by M_1^{-1} corresponds to a series of row operations on $(M_1 \mid M_2)$, which means that the above row labels are elements of the ideal $\langle f, g\rangle$. In particular, the polynomials $\underline{x^2y} - y, \underline{xy^2} - x, \underline{x^2} - 1, \underline{y^2} - 1$ are in $\langle f, g\rangle$.

The columns in the above matrices are labeled by monomials of degree ≤ 3. Plotting their exponent vectors gives the picture



where B is shown in black and the underlined monomials give the boxes. Note that the boxes line on the *border* of B—they are *border monomials*, and the four polynomials $\underline{x^2y} - y, \underline{xy^2} - x, \underline{x^2} - 1, \underline{y^2} - 1$ form a *border basis* of $\langle f, g\rangle$. You can read more about border bases in [**232**].

Once we have the border basis, the multiplication matrices

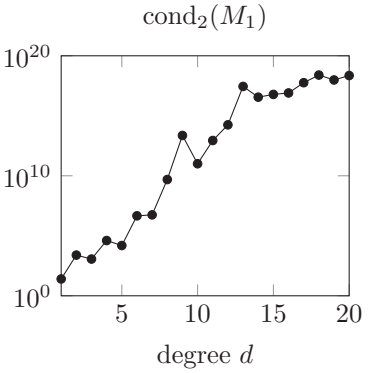
$$M_x = \begin{array}{c} \\ 1 \\ x \\ y \\ xy \end{array} \begin{array}{c} \begin{array}{cccc} x\cdot 1 & \underline{x\cdot x} & x\cdot y & \underline{x\cdot xy} \end{array} \\ \left[\begin{array}{cccc} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right] \end{array}, \quad M_y = \begin{array}{c} \\ 1 \\ x \\ y \\ xy \end{array} \begin{array}{c} \begin{array}{cccc} y\cdot 1 & y\cdot x & \underline{y\cdot y} & \underline{y\cdot xy} \end{array} \\ \left[\begin{array}{cccc} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right] \end{array}$$

are easy to find. When we multiply B by x, we get either an element of B or a border monomial. For the former, the column of M_x is obvious, and for the latter, the border basis gives the column. The same thing happens with M_y, so you can see the close link between border bases and the multiplication matrices. ◁▷

The same happens for generic f, g of degree d. In this case, the Macaulay matrix $M = (M_1 \mid M_2)$ has size $2\binom{d+1}{2} \times \binom{2d+1}{2} = d(d+1) \times d(2d+1)$, where M_1 is $d(d+1)\times d(d+1)$ and invertible. The process described in Example 2.9 generalizes, where appropriate rows of $(I_{d(d+1)} \mid M_1^{-1}M_2)$ give a border basis, from which M_x and M_y are easily computed.

The upshot is that we have a lovely method to compute M_x and M_y. But while the theory is wonderful, the reality is that the above process can be ill-conditioned numerically. Looking back at Example 2.9, you can see that the key step is multiplying by M_1^{-1}. Hence, in order to know accurately we have computed M_x and M_y, we need to know the condition number $\mathrm{cond}_2(M_1)$.

In [**343**], Telen and Van Barel compute the condition number $\mathrm{cond}_2(M_1)$ for random f, g of degree d with $1 \le d \le 20$:

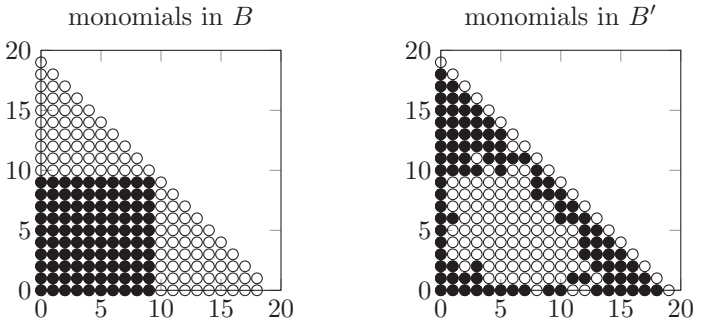


These are poorly conditioned matrices!

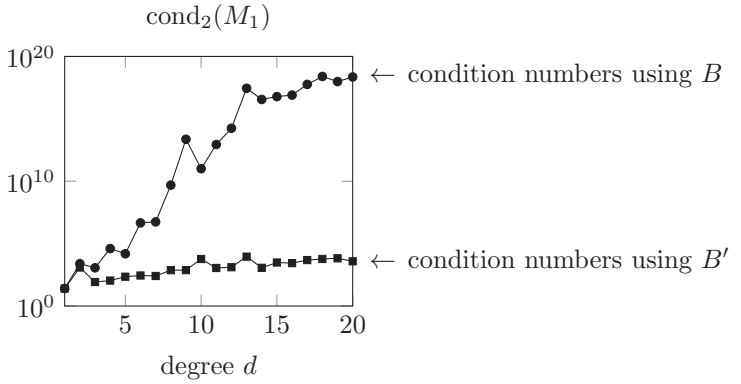
This is where numerical linear algebra comes to the rescue. The hope is that switching from B to a different basis B' of $\mathbb{C}[x,y]/\langle f, g\rangle$ will lead to smaller condition numbers. The paper [**343**] shows how to find a monomial basis B' with this property. Their strategy is to do numerical linear algebra on $M = (M_1 \,|\, M_2)$ *before* inverting M_1 by using QR factorization with optimal column pivots. Since columns correspond to monomials, switching columns gives a different monomial basis B'.

The full details are given in [**343**, Algorithm 1]. Here is an example with $d = 10$ that shows the original monomial basis B and the new basis B':

(2.4)



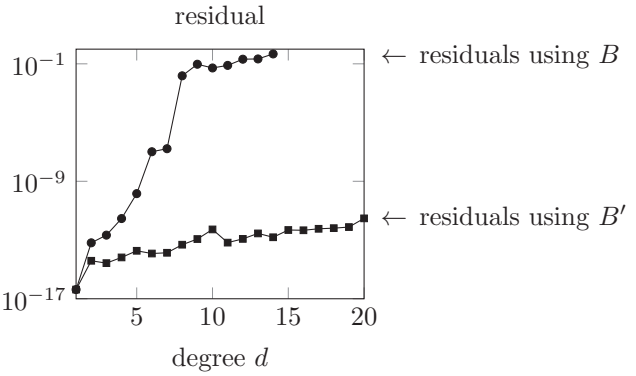
On the left, we see the nice order ideal B, while the monomials in B' are more scattered—neither an order ideal nor connected to 1. But the condition numbers are much better:



It is clear that from a numerical point of view, one wants to compute multiplication matrices using B', not B.

Another way to measure the improvement is to compute residuals. This works as follows. Simultaneously diagonalize M_x, M_y to get approximate solutions (x_i, y_i), and normalize f and g so that the vector norm of their coefficients is 1. Then the *residual* is the maximum of $\frac{1}{2}(|f(x_i, y_i)|+|g(x_i, y_i)|)$ over all solutions. Here, smaller residuals are better.

In [**343**], residuals are computed for random choices of f, g with $1 \leq d \leq 20$, using both B and B'. The improvement is dramatic:

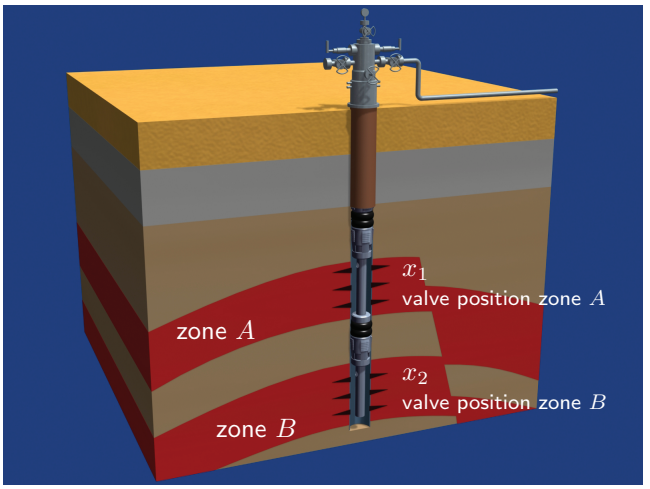


The paper [**343**] focuses on the dense case, while its follow-up [**344**] by Telen, Mourrain and Van Barel extends the analysis to sparse polynomials.

Algebraic Oil. Our final example involves a collaboration between Shell International and the Universities of Genoa and Passau. The goal is to model in a well using algebraic geometry and numerical linear algebra. Papers written on this subject include [**199**] and [**242**]. The authors of [**242**] are Kreuzer (Passau), Poulisse (Shell) and Robbiano (Genoa).

For concreteness, we will focus on an example from [**13**] by Baciu and Kreuzer that involves the two-zone well shown here:

(2.5)



Although the picture looks nice, we actually know little about the underground geometry beyond the fact that there are two zones. So it's not possible to build a model of oil production in advance—one needs to create the model based on observations.

In the picture (2.5), x_1 and x_2 record the settings of the valves at A and B. Altogether, there are seven variables in this model:

$$\begin{aligned} x_1 &= v_A = \text{valve opening at } A \\ x_2 &= v_B = \text{valve opening at } B \\ x_3 &= p_A = \text{pressure drop across valve } A \\ x_4 &= p_B = \text{pressure drop across valve } B \\ x_5 &= p_{AB} = \text{pressure loss between } A \text{ and } B \\ x_6 &= p_{tr} = \text{pressure loss } A \text{ to the surface} \\ x_7 &= p_{gas} = \text{gas production (related to another pressure)} \end{aligned} \tag{2.6}$$

We want to find a set of monomials in $x_1, \dots, x_7$ such that

$$\text{oil production} \approx c_1 x^{\alpha_1} + \cdots + c_s x^{\alpha_s} \tag{2.7}$$

for constants $c_1, \dots, c_s$. We like monomials because they can represent interactions between variables. For example, $x_1 x_4 = v_A p_B$ measures how the setting of value A mediates the influence on oil production of the pressure drop across valve B.

In this setting, a data point consists of $(p, P) \in \mathbb{R}^7 \times \mathbb{R}$, where p records the values of $x_1, \dots, x_7$ and $P \in \mathbb{R}$ is the corresponding oil production for these values. There are 5500 data points, giving

$$\mathcal{X} = \{p_1, \dots, p_{5500}\} \subseteq \mathbb{R}^7$$

and oil production

$$\mathcal{P} = (P_1, \dots, P_{5500}) \in \mathbb{R}^{5500}.$$

A Naive Model. For the moment, assume that our data is exact and that we can do exact computations. The *vanishing ideal* of $\mathcal{X}$ is

$$\mathbf{I}(\mathcal{X}) = \{f \in \mathbb{R}[x_1, \dots, x_7] \mid f(p_i) = 0,\ i = 1, \dots, 5500\}.$$

Then evaluation at $\mathcal{X} = \{p_1, \dots, p_{5500}\}$ induces an $\mathbb{R}$-algebra isomorphism

$$\mathbb{R}[x_1, \dots, x_7]/\mathbf{I}(\mathcal{X}) \xrightarrow{\text{eval}_{\mathcal{X}}} \mathbb{R}^{5500}.$$

If B is a monomial basis of $\mathbb{R}[x_1, \dots, x_7]/\mathbf{I}(\mathcal{X})$, then $\text{eval}_{\mathcal{X}}(B)$ is a basis of $\mathbb{R}^{5500}$. Expressing the production data $\mathcal{P} = (P_1, \dots, P_{5500})$ in terms of this basis gives

$$\mathcal{P} = \sum_{x^\alpha \in B} c_\alpha \text{eval}_{\mathcal{X}}(x^\alpha).$$

Then the polynomial $F = \sum_{x^\alpha \in B} c_\alpha x^\alpha$ satisfies (2.7) perfectly.

From the point of view of interpolation, this works nicely since $F(p_i) = P_i$ for all i. But from a modeling perspective, it is terrible, since the model is overly complex (5500 monomials!) and suffers from the classic problem of over-fitting. In general, one wants to avoid situations where "the model is the data."

A better way to proceed is to pick a subset $B_0 \subseteq B$ so that B_0 is reasonably small and $\mathcal{P} \in \mathbb{R}^{5500}$ is reasonably close to the subspace $\text{Span}(\text{eval}_{\mathcal{X}}(B_0)) \subseteq \mathbb{R}^{5500}$. Let $\mathcal{P}_0$ denote the orthogonal projection of $\mathcal{P}$ onto this subspace. Then writing $\mathcal{P}_0 = \sum_{x^\alpha \in B_0} c_\alpha \text{eval}_{\mathcal{X}}(x^\alpha)$ gives $F_0 = \sum_{x^\alpha \in B_0} c_\alpha x^\alpha$ that satisfies (2.7).

But how to do pick B_0? The first step is to understand where the monomial basis B comes from.

The Buchberger-Möller Algorithm. Since the evaluation map

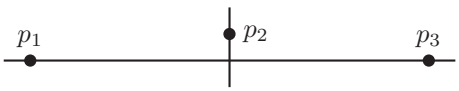
$$\text{eval}_{\mathcal{X}} : \mathbb{R}[x_1, \ldots, x_7] \longrightarrow \mathbb{R}^{5500}$$

is onto, we find a monomial basis of $\mathbb{R}[x_1, \ldots, x_7]/\mathbf{I}(\mathcal{X})$ by listing all monomials $x^{\alpha_1}, x^{\alpha_2}, \ldots$, successively applying $\text{eval}_{\mathcal{X}}$, and discarding those whose images are linearly dependent on the images of the previous ones. Once we have found 5500 monomials, we have a monomial basis B.

A better approach is due to Buchberger and Möller [**49**] in 1982. Monomials are ordered using a monomial order, and keeping track of linear relations coming from discarded monomials leads to a Gröbner basis G for $\mathbf{I}(\mathcal{X})$ such that B is the order ideal consisting of all remainder monomials for division by G. A description of how this works for lex order can be found in [**105**, Section 2.3].

The numerical instability of the Buchberger-Möller algorithm is well known. Here is a simple example,

EXAMPLE 2.10. Let $\mathcal{X} = \{p_1, p_2, p_3\} = \{(-1, 0), (0, c), (1, 0)\} \subseteq \mathbb{R}^2$, where c is a parameter, possibly 0.



Using graded lex with $x > y$, the Buchberger-Möller algorithm gives the following reduced Gröbner bases G for $\mathbf{I}(\mathcal{X})$ and monomial bases B for $\mathbb{R}[x, y]/\mathbf{I}(\mathcal{X})$:

$$c = 0\text{: } G_0 = \left\{x^3 - x, y\right\},\ B_0 = \left\{1, x, x^2\right\}$$
$$c \neq 0\text{: } G_c = \left\{x^2 - 1 - \tfrac{1}{c}y, xy, y^2 - cy\right\},\ B_c = \left\{1, y, x\right\}.$$

When c is large, G_c and B_c are what we want. But when c is not 0 but really close to 0, then we would prefer G_0 and B_0. One way to achieve this is to replace "exactly linearly dependent" with "approximately linearly dependent" when considering the evaluation of a new monomial.

To see what this means, specify a tolerance ε for approximate linear dependence, and suppose $0 < c < \varepsilon$. Listed in increasing order for graded lex with $x > y$, the monomials are $1, y, x, y^2, xy, x^2, \ldots$. When we evaluate these at $\mathcal{X}$ and consider linear dependence, we get the following table:

x^α	$\text{eval}_{\mathcal{X}}(x^\alpha)$	B_{exact}	B_{approx}
1	$(1, 1, 1)$	$\{1\}$	$\{1\}$
y	$(0, c, 0)$	$\{1, y\}$	$\{1\}$
x	$(-1, 0, 1)$	$\{1, y, x\}$	$\{1, x\}$
y^2	$(0, c^2, 0)$	$\{1, y, x\}$	$\{1, x\}$
xy	$(0, 0, 0)$	$\{1, y, x\}$	$\{1, x\}$
x^2	$(1, 0, 1)$	$\{1, y, x\}$	$\{1, x, x^2\}$

The line for y is revealing, since $(0, c, 0)$ and $(1, 1, 1)$ are linearly independent but ε-linearly dependent. This explains why y is added to B_{exact} but not B_{approx}. The other entries in the table are similar. Since $B_{\text{approx}} = B_0$, we are happy. ◁▷

In the above example, we did not say exactly what we mean by ε-linearly dependent. Making this precise takes careful thought and can be done in several

ways. The paper [**199**] introduces the *Approximate Vanishing Ideal algorithm*, and the 2014 PhD thesis of Limbeck [**257**] studies this algorithm in detail and describes a variant called the *Approximate Buchberger-Möller algorithm*. For expository purposes, we will say a few words about the latter.

The Approximate Buchberger-Möller Algorithm. We begin with a modified Buchberger-Möller algorithm that focuses on the monomial basis B. Given points $\mathcal{X} = \{p_1, \ldots, p_s\} \subseteq \mathbb{R}^n$, we have an isomorphism

$$\mathbb{R}[x_1, \ldots, x_n]/\mathbf{I}(\mathcal{X}) \xrightarrow{\text{eval}_{\mathcal{X}}} \mathbb{R}^s.$$

Pick a graded monomial order $<$ on $\mathbb{R}[x_1, \ldots, x_n]$ and list monomials in increasing order according to $<$. We add monomials one at a time to B, starting with $B = \{1\}$. After processing the monomials of degree d, consider the monomials

$$L = \{\text{monomials of degree } d+1\} \cap \big(x_1 B \cup \cdots \cup x_n B\big). \tag{2.8}$$

At this point, B consists of monomials of degree $\leq d$, so the important monomials of degree $d+1$ to consider are those on the border of B. This is the set L defined above. We work through the elements of L according to $<$. For each $x^\alpha \in L$, we compute $\text{eval}_{\mathcal{X}}(x^\alpha)$. Then:

- If $\text{eval}_{\mathcal{X}}(x^\alpha) \cup \text{eval}_{\mathcal{X}}(x^\alpha)(B)$ is linearly dependent, then discard x^α.
- Otherwise, add x^α to B.

Once all monomials in L have been processed, we move to the next degree and continue until $L = \emptyset$. By [**257**, Algorithm 22 with $\varepsilon = 0$ and Theorem 4.3.1], the algorithm terminates. This is the *exact algorithm*, which has two nice properties:

- The final value of B is an order ideal basis of $\mathbb{R}[x_1, \ldots, x_n]/\mathbf{I}(\mathcal{X})$.
- By keeping track of the linear dependencies of discarded monomials, one can obtain a border basis for $\mathbf{I}(\mathcal{X})$.

For the approximate version of this algorithm, we follow [**257**, Algorithm 22]. Fix a threshold $\varepsilon > 0$. As in the exact algorithm, suppose that we are deciding whether to add x^α to B. Consider the $(1 + |B|) \times s$ matrix A whose first column is $\text{eval}_{\mathcal{X}}(x^\alpha)$ and remaining columns are $\text{eval}_{\mathcal{X}}(x^\beta)$ for $x^\beta \in B$ in decreasing order according to $>$. (We will give an example of this below.) Then $A^T A$ is a symmetric semi-definite matrix whose entries the dot products of the columns of A. Its smallest eigenvalue λ has the property that

$$\sqrt{\lambda} = \min_{\|v\|_2 = 1} \|Av\|_2. \tag{2.9}$$

There are two methods to find λ: compute a SVD of $A^T A$ or solve the above minimization problem by numerical methods. Now we decide what happens to x^α as follows:

- If $\sqrt{\lambda} \leq \varepsilon$, then discard x^α.
- Otherwise, add x^α to B.

Once the monomials of degree d are processed in this way, we define L as in (2.8) and continue until $L = \emptyset$. We call this the *ε-algorithm*. By [**257**, Theorem 4.3.1]:

- B is connected to 1, though it may fail to be an order ideal.
- If B is an order ideal, then when $\sqrt{\lambda} \leq \varepsilon$, eigenvectors for λ serve the role of ε-linear dependencies, and from these, one can obtain a set G of polynomials that generate an *approximate vanishing ideal* of $\mathcal{X}$.

The full theory of these concepts is explained in [**199**] and [**257**]; a nice exposition appears in [**242**]. For an algebraist, an approximate vanishing ideal is troubling since its generators usually generate the unit ideal in the standard sense.

Ignoring these complications, there is one elementary feature of the ε-algorithm we need to mention. By (2.9),

$$\begin{aligned}\lambda = 0 &\iff \text{columns of } A \text{ are linearly dependent}\\ &\iff \operatorname{eval}_{\mathcal{X}}(x^\alpha) \cup \operatorname{eval}_{\mathcal{X}}(x^\alpha)(B) \text{ is linearly dependent.}\end{aligned}$$

When $\lambda = 0$, we have $\sqrt{\lambda} = 0 \le \varepsilon$, so any monomial discarded by the exact version of the algorithm will also be discarded by the ε-algorithm. However, the condition $\sqrt{\lambda} \le \varepsilon$ gives us leeway to discard more monomials. The result is that the set B produced by the ε-algorithm may fail to be a basis of $\mathbb{R}[x_1, \dots, x_n]/\mathbf{I}(\mathcal{X})$.

Here is an example due to Limbeck [**257**, Example 4.3.8].

EXAMPLE 2.11. Let $\mathcal{X} = \{(0,0), (0,1), (1,0), (1,1), (.5,.5)\} \subseteq \mathbb{R}^2$ consist of the vertices of the unit square and its center. We will use graded lex order with $x > y$ and tolerance $\varepsilon = 0.2$. The ε-algorithm starts with $B = \{1\}$ and successively adds x and then y to obtain $B = \{y, x, 1\}$. At this point, L is defined to be $\{x^2, xy, y^2\}$, and to decide about x^2, the matrix A is

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ .25 & .5 & .5 & 1 \end{bmatrix}.$$

The smallest eigenvalue of A^TA is $\lambda = .0246396$, and $\sqrt{\lambda} = .15697 \le \varepsilon$. Hence we discard x^2. When we turn to xy, we get a new first column, so A becomes

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ .25 & .5 & .5 & 1 \end{bmatrix}.$$

The smallest eigenvalue of A^TA is now $\lambda = .145979$, and $\sqrt{\lambda} = .38207 > \varepsilon$. Thus we add xy to obtain $B = \{xy, y, x, 1\}$.

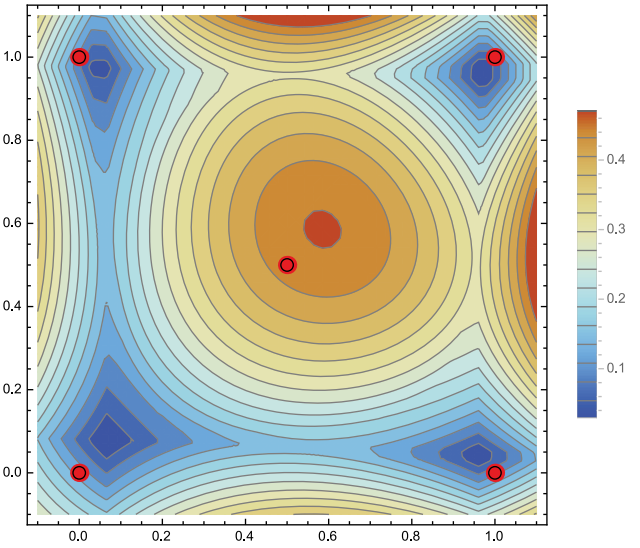
One can check that y^2 is also discarded, so we are done with degree 2. Then we define $L = \{x^2y, xy^2\}$ (the degree 3 elements of $xB \cup yB$), and both of these are discarded as well. So B is still $\{xy, y, x, 1\}$. It follows that when we increase the degree to 4, we get $L = \emptyset$ and terminate.

There are a lot of interesting things going on here. The set $B = \{xy, y, x, 1\}$ has four elements, yet $\dim \mathbb{R}[x,y]/\mathbf{I}(\mathcal{X}) = 5$ since we have five points. So B is *not* a basis of $\mathbb{R}[x,y]/\mathbf{I}(\mathcal{X})$. Also, the four discarded monomials are x^2, y^2, x^2y, xy^2. Using eigenvectors of λ to represent ε-linear dependence, one gets polynomials

$$\begin{aligned} g_1 &= -.697x^2 + .715x - .044\\ g_2 &= .685y^2 + .041xy - .724y - .021x + .054\\ g_3 &= -.698x^2y + .715xy - .008y - .008x - .013\\ g_4 &= -.698xy^2 + .715xy - .008y - .008x - .013 \end{aligned} \tag{2.10}$$

that generate the approximate vanishing ideal of $\mathcal{X}$.

The polynomials g_1, g_2, g_3, g_4 generate the unit ideal in $\mathbb{R}[x, y]$, and if one asks a program such as MATHEMATICA to solve $g_1 = g_2 = g_3 = g_4 = 0$, one gets the empty set. To get a better sense of how the g_i relate to the five points in $\mathcal{X}$, here is a picture that shows $\mathcal{X}$ in red along with a contour plot of $|g_1| + |g_2| + |g_3| + |g_4|$:



You can see that the g_i are not very small at the central red point but get close to zero in the dark blue areas. So we lose the center of the square, but its influence is felt since the dark blue areas have been pulled towards the center.

This example makes it clear that the approximate world requires us to think differently about basic objects in algebra and geometry. ◁▷

One exciting feature of the ε-algorithm is that it can produce a set B with fewer monomials than the number of points in $\mathcal{X}$. In a sense, the algorithm is saying that because of the tolerance ε, fewer monomials may be required.

In Example 2.11, we started with five points but ended up with only four monomials. In some cases, the reduction is more dramatic, as we will see when we apply these ideas to the oil production model discussed earlier in the section.

Back to the Oil Well. First observe that when the valve settings x_1, x_2 in (2.5) are both zero, there is no oil production. Hence the oil production polynomial we seek should lie in the ideal $\langle x_1, x_2 \rangle \subseteq \mathbb{R}[x_1, x_2]$. This restriction leads to the *subideal Buchberger-Möller algorithm*, which computes a monomial basis of

$$\langle x_1, x_2 \rangle / (\mathbf{I}(\mathcal{X}) \cap \langle x_1, x_2 \rangle) \longhookrightarrow \mathbb{R}[x_1, \ldots, x_7]/\mathbf{I}(\mathcal{X}).$$

This algorithm and its approximate variants are developed in [**240**].

In [**13**], an approximate subideal algorithm is applied to the 5500 data points $\mathcal{X} = \{p_1, \ldots, p_{5500}\} \subseteq \mathbb{R}^7$. The result is a set B_0 consisting of

31 monomials

in the variables $x_1, \ldots, x_7$ from (2.6). When we orthogonally project the oil production data $\mathcal{P} = \{P_1, \ldots, P_{5500}\} \in \mathbb{R}^{5500}$ onto the span of $\mathrm{eval}_{\mathcal{X}}(B_0)$, we get a

polynomial with 31 terms

$$F_0 = \sum_{x^\alpha \in B_0} c_\alpha x^\alpha \tag{2.11}$$

that models oil production. The precise formula for F_0 is given in [**13**], along with a discussion of how well the model fits the data.

An important observation is that in this model, the pressure variables $x_3, \dots, x_7$ in (2.6) serve as proxies for basic physical quantities such as kinetic energy, work, and dissipation that cannot be measured directly. Unfortunately, many of the 31 monomials in the model don't make sense physically, and the same is true for the polynomials g_1, g_2, g_3, g_4. Creating an algorithm that respects the physics without knowing it in advance is a continuing challenge.

A different analysis of this example that focuses on the five pressure variables can be found in [**242**].

Final Comments. In this section, we explored two major examples that used a variety of numerical methods from linear algebra:

- To get well-conditioned multiplication matrices for generic polynomials in $\mathbb{C}[x, y]$ of degree d, the paper [**343**] used QR factorization with column pivots. In (2.4), the nice order ideal B was replaced with the more scattered B' that gave multiplication matrices with much better condition numbers.
- To model oil production, the paper [**13**] used an ε-approximate version of linear dependence to create a small set of monomials that gave the desired model (2.11). But as we saw in Example 2.11, these methods replace the vanishing ideal with polynomials that usually generate the unit ideal.

In both examples, we worked in a finite-dimensional quotient of a polynomial ring by an ideal. In the first example, we knew the ideal, but to get the solutions via multiplication matrices, we had to abandon working with an order ideal of monomials. In the second example, we had the solutions (the data points), but to get a reasonable set of monomials, we had to abandon the usual notion of ideal.

The moral is that while numerical linear algebra has a lot to offer algebraic geometry, one needs to approach the whole enterprise with an open mind and a willingness to think outside the box.

2.2. Homotopy Continuation and Applications

The previous section harnessed the power of numerical linear algebra to help study systems of polynomial equations. In this section, we will tap into the equally rich collection of numerical techniques used to solve ordinary differential equations. These techniques are relevant to algebraic geometry because *homotopy continuation* enables us to solve a polynomial system by tracking solutions of a suitable system of ODEs.

The plan is to explain how homotopy continuation works and then explore four substantial examples.

The Idea of Homotopy Continuation. Suppose that we want to compute all complex solutions of a polynomial system $f_1(z) = \cdots = f_n(z) = 0$ for $f_i \in \mathbb{C}[z] = \mathbb{C}[z_1, \dots, z_n]$. Note that the number of equations matches the number of variables. For the time being, we will assume that there are finitely many solutions. We also

write the variables as $z = (z_1, \ldots, z_n)$ to emphasize that we are working over the complex numbers.

In homotopy continuation, we have the *target system* $F(z) = (f_1(z), \ldots, f_n(z))$ together with a *start system* $G(z) = (g_1(z), \ldots, g_n(z))$ whose solutions we know. We get from one system to the other via the homotopy

$$H(z,t) = tG(z) + (1-t)F(z), \quad t \in [0,1]. \tag{2.12}$$

Thus $H(z,1)$ is the start system $G(z)$ and $H(z,0)$ is the target system $F(z)$. The target system is $t = 0$ since there are more floating point numbers near 0.

The basic idea is that as we decrease t from 1 to 0, $H(z,t)$ deforms the start system $G(z) = 0$ into the target system $F(z) = 0$. In the process, we hope that solutions of $G(z) = 0$ will deform into solutions of $F(z) = 0$.

To formulate this precisely, let z_0 be a solution of $G(z) = 0$ and assume that $z(t) : [0,1] \to \mathbb{C}^n$ is a differentiable function satisfying $z(1) = z_0$. By the chain rule,

$$\frac{d}{dt}H(z(t),t) = \sum_{i=1}^{n} \frac{\partial H}{\partial z_i}\big(z(t),t\big)\, z_i'(t) + \frac{\partial H}{\partial t}\big(z(t),t\big).$$

Hence the following are equivalent:

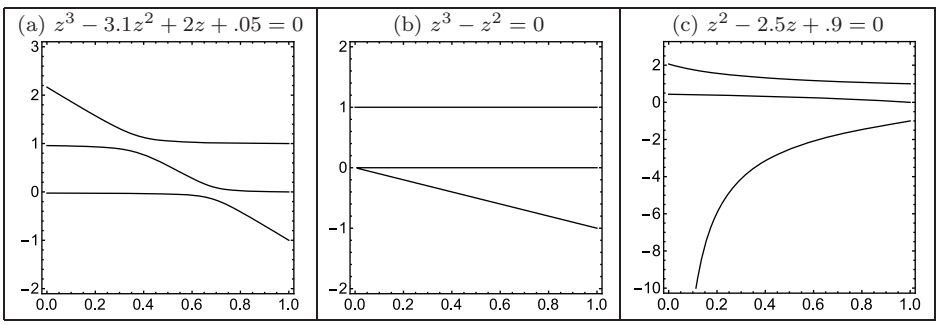
- $H(z(t),t) = 0$ for all t and $z(1) = z_0$.
- $z(t)$ satisfies the initial value problem (IVP)

$$\sum_{i=1}^{n} \frac{\partial H}{\partial z_i}\big(x(t),t\big)\, z_i'(t) + \frac{\partial H}{\partial t}\big(z(t),t\big) = 0, \quad z(1) = z_0. \tag{2.13}$$

A solution $z(t)$ of (2.13) is called *path*. If the path is defined on all of $[0,1]$, then $z(t)$ deforms the solution $z(1) = z_0$ of the start system $G(z) = 0$ into a solution $z(0)$ of the target system $F(z) = 0$, exactly as we had hoped.

However, not all solutions of the start system deform so nicely. In fact, there are many possible behaviors for a given path. Let's begin with three good behaviors.

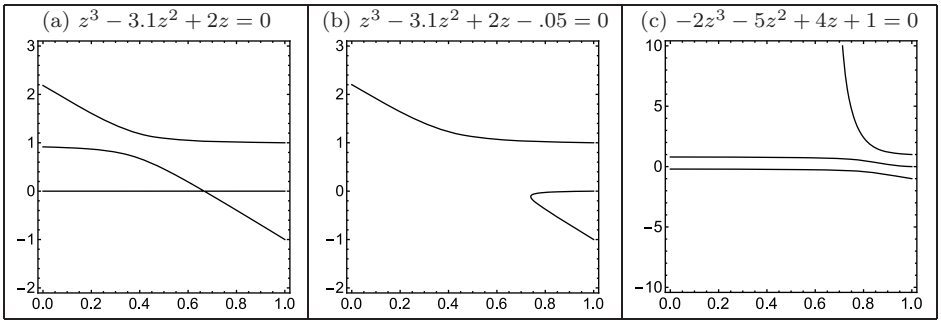
EXAMPLE 2.12. We will use the start system $z^3 - z = z(z^2 - 1) = 0$ with solutions $z = 0, \pm 1$. Here are the paths we get for the three different target systems:



- In (a), the paths go from $0, \pm 1$ to the roots of $z^3 - 3.1z^2 + 2z + .05 = 0$. This is as good as it gets.
- In (b), two paths meet when $t = 0$, which reflects the fact that 0 is a double root of $z^3 - z^2 = 0$.
- In (c), one of the paths goes to ∞, which makes sense since the target system is the quadratic $z^2 - 2.5z + .9 = 0$.

It may seem odd to claim that (b) and (c) represent "good" behavior. But in practice, it happens often that start system has more solutions than the target system, so that some solutions must either coincide or go to ∞ as $t \to 0^+$. ◁▷

EXAMPLE 2.13. With the same start system, we now give examples of some bad behaviors that can occur:



- In (a), two paths cross when $t = 2/3$, which means that the IVP fails to have a unique solution. This makes it harder to follow paths numerically.
- In (b), as we go from 1 to 0, two paths meet as $t \to 0.73970^+$, and neither can be continued over $\mathbb{R}$.
- In (c), the deformed system drops degree when $t = 2/3$, which means that one path goes to ∞ as $t \to 2/3^+$.

These problems arise because the deformed system fails to be generic for all $t > 0$. In (a) and (b), $H(z,t)$ acquires a double root, while in (c), it drops degree. ◁▷

We can avoid the problems encountered in Example 2.13 by working over $\mathbb{C}$ and using the *gamma trick*. To explain this, note that the homotopy (2.12) is a line segment connecting $G(z)$ to $F(z)$. Bad things happen when the line segment intersects loci where the degree drops or multiple roots exist. These bad loci are defined by algebraic conditions, which over $\mathbb{C}$ have real codimension at least two. Since a line segment has dimension 1, we have a good chance of missing the bad loci provided we wiggle the line segment a bit.

The gamma trick achieves this by replacing the line segment $[0,1]$ with the circular arc in the plane defined by

$$\tau(t) = \frac{\gamma t}{\gamma t + (1-t)}, \quad t \in [0,1].$$

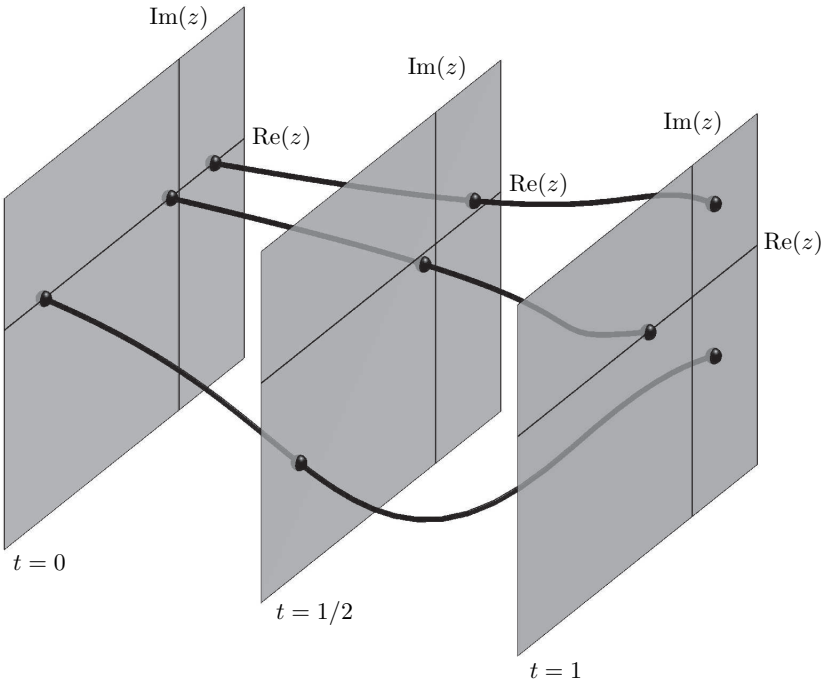
where we assume $\gamma \in \mathbb{C} \setminus \mathbb{R}$ (to ensure that the denominator never vanishes) and $|\gamma| = 1$ (for numerical reasons). Since $\tau(0) = 0$ and $\tau(1) = 1$, we get the new homotopy

$$H(z, \tau(t)) = \tau(t)G(z) + (1 - \tau(t))F(z) \tag{2.14}$$

that connects $F(z)$ (when $t = 0$) to $G(z)$ (when $t = 1$). Lemma 7.13 of [**324**] ensures that with finitely many exceptions, this homotopy will avoid the bad loci.

Another advantage of working over $\mathbb{C}$ is that for the target system $-2z^3 - 5z^2 + 4z + 1 = 0$ from (c) in Example 2.13, we can replace the start system $z^3 - z = 0$ with $z^3 + 1 = 0$, whose roots are the negative cube roots of unity. This is good from the numerical point of view since they are equally spaced around the unit circle.

Then picking $\gamma = .4 + .77i$ gives the following splendid picture from [**21**, p. 26]:



Making Homotopy Continuation Work. We know that with high probability, a randomly chosen γ with $|\gamma| = 1$ will avoid the bad loci. But there are many other issues that need to be addressed in order for homotopy continuation to be a useful tool for solving equations:

- Given the target system $F(z) = 0$, explain how to pick a suitable start system $G(z) = 0$.
- Prove that all *isolated solutions* of $F(z) = 0$ are found in this way.
- Decide if a path goes to ∞.
- Understand the *endgame* as $t \to 0^+$.
- Understand *condition numbers* and *adaptive precision*.
- Decide if two paths give the same solution (*multiplicity*).
- Handle *positive dimensional* components of the solution variety.
- Handle systems where the number of variables doesn't equal the number of equations.

One could write a book about this, and in fact there are two books [**21**] and [**324**] that go into great detail about the above issues. There is also the book *Condition* [**50**] devoted to the study of condition numbers.

In addition to a firm theoretical foundation, we are fortunate to have excellent software that implements homotopy continuation:

- Bertini [**20**], named for the Bertini Theorems.
- PHCpack [**365**], for Polyhedral Homotopy Continuation package.
- NAGM2 [**258**], for Numerical Algebraic Geometry for Macaulay2.
- HOM4PS [**78**], for Homotopy for Polynomial Systems.

Higher Dimensional Components. The solution set of a system of equations breaks up into irreducible components of various dimensions. So far, we have assumed that each component is a point (the zero dimensional case), which can be represented numerically by an approximation of the point and found (for example) by homotopy continuation.

Now suppose that an irreducible component V of the solution set has positive dimension. How do we find V numerically? How do we represent V numerically? One way would be compute some sort of "generic point" of V. In Chapter 1, we discussed several notions of generic point. Suppose that $V \subseteq \mathbb{C}^n$ is defined by a prime ideal $\mathfrak{p} \subseteq R = \mathbb{C}[x_1, \ldots, x_n]$. Recall from Section 1.2 that we encountered three types of generic point:

Type	Point	Description
Enriques & Chisini	$p \in V \setminus W$	W = subvariety of points to avoid
van der Waerden	$(\xi_1, \ldots, \xi_n)$	$\xi_i = x_i + \mathfrak{p} \in R/\mathfrak{p} = \mathbb{C}[V] \subseteq \mathbb{C}(V)$
Grothendieck	$\mathfrak{p} \in \operatorname{Spec}(R)$	$V = \mathbf{V}(\mathfrak{p})$, $\mathfrak{p}$ prime ideal

The last two entries in the table give full information about V but can't be done numerically. The first makes sense numerically (compute a random point of V), but a single numerical point gives little information about V. One exception is $V = \mathbb{C}^n$, where a random point is useful—for n variables, make a random choice for each variable.

To apply this to an arbitrary irreducible variety $V \subseteq \mathbb{C}^n$ of dimension k and degree d, we will use the algebra and geometry of *Noether normalization*.

Noether Normalization. For V as above, Noether normalization (see, for example, [**104**, Chapter 5, §6]) tells us that if we pick generic linear forms $L_1, \ldots, L_k$ in $x_1, \ldots, x_n$, then $\mathbb{C}[L_1, \ldots, L_k] \subseteq R = \mathbb{C}[x_1, \ldots, x_n]$ induces an inclusion

$$\mathbb{C}[L_1, \ldots, L_k] \hookrightarrow R/\mathfrak{p} = \mathbb{C}[V] \tag{2.15}$$

such that $\mathbb{C}[V]$ is integral over $\mathbb{C}[L_1, \ldots, L_k]$. Here, $R = \mathbb{C}[x_1, \ldots, x_n]$ and $\mathfrak{p} = \mathbf{I}(V)$, and we write (2.15) as $\mathbb{C}[L_1, \ldots, L_k] \subseteq \mathbb{C}[V]$. Geometrically, this inclusion is the map on coordinate rings induced by the linear map

$$\pi : V \longrightarrow \mathbb{C}^k, \quad \pi(p) = (L_1(p), \ldots, L_k(p)). \tag{2.16}$$

Furthermore, (2.15) being an integral extension of rings tells us that (2.16) is a finite morphism of varieties.

Monodromy. The inclusion (2.15) and the morphism (2.16) lead to interesting groups as follows. Recall that V has dimension k and degree d. Then:

Algebra: We get a field extension $\mathbb{C}(L_1, \ldots, L_k) \subseteq \mathbb{C}(V)$ of degree d. The minimal polynomial of a primitive element of this extension has the *Galois group* $G \subseteq S_d$.	*Geometry*: We get a finite fiber $\pi^{-1}(q) = \{p_1, \ldots, p_d\}$ for $q \in \mathbb{C}^k$ random. A generic loop γ at q lifts to a path in V with $\hat{\gamma}(0) = p_i$, $\hat{\gamma}(1) = p_{\sigma(i)}$, $\sigma \in S_d$. This gives the *monodromy group* $G \subseteq S_d$.

The wonderful fact is that these groups are equal up to conjugation, as proved by Harris in 1979 [**187**]. Since the Galois group of an irreducible polynomial acts transitively on its roots, we see that the monodromy group acts transitively on the points of the fiber $\pi^{-1}(q)$.

Path lifting can be done numerically, which gives a numerical method to compute the monodromy group.

Witness Point Sets. For our irreducible component $V \subseteq \mathbb{C}^n$ of dimension k and degree d, random linear forms $L_1, \ldots, L_k$ give a finite morphism $\pi : V \to \mathbb{C}^k$ of degree d. For a random $q \in \mathbb{C}^k$, we have

$$\pi^{-1}(q) = V \cap L, \ L = \text{random affine subspace of codimension } k \text{ in } \mathbb{C}^n.$$

The set $V \cap L$ has d points and comes equipped with a $G \subseteq S_d$. The points of $V \cap L$ and the group G can all be computed numerically.

We call $V \cap L$ a *witness point set* of V. This is the numerical replacement for a generic point of V. Unlike the algebraic setting of van der Waerden or Grothendieck, a witness point set doesn't directly allow us to recover V but still contains a lot of useful information, including dimension, degree, and monodromy. Furthermore, by computing multiple witness point sets (this is the idea of *sampling*), one can sometimes build up a vivid picture of V. You will learn more about sampling in Section 2.3.

However, our discussion so far is unrealistic, since in practice we don't know the irreducible components of the solution set in advance. So now suppose that V is the solution set of a system of polynomials. The goal is now to find a witness point set of each irreducible component of V. This is what is called *numerical irreducible decomposition*.

Computing a Numerical Irreducible Decomposition. Given a system of equations, the dimension of the solution set $V \subseteq \mathbb{C}^n$ is the largest integer k such that $V \cap L \neq \emptyset$ for a random affine subspace $L \subseteq \mathbb{C}^n$ of codimension k. This can be determined numerically, so we can assume that we know the dimension k of V.

Noether normalization still applies in this situation, so that if L is defined by $L_1 = \cdots = L_k = 0$, then we get a finite morphism $\pi : V \to \mathbb{C}^k$ as in (2.16), giving a candidate witness point set $\pi^{-1}(q) = V \cap L$ as above. Since L is generic of codimension k, it will miss all components of V of dimension strictly smaller than k. It follows that $V \cap L$ is *a union of witness point sets of the dimension k components of V*.

We still have a monodromy group G in this situation, though it may fail to act transitively on $V \cap L$. The key observation is that the *orbits* of G on $V \cap L$ correspond to the irreducible components of V of dimension k. Furthermore, each orbit is a witness point set of the corresponding component, and the number of elements in the orbit is the degree of the component. Hence we have witness point sets for the top-dimensional irreducible components of V. In practice, several tools are used to achieve this decomposition, not just monodromy. See [**21**, Chapter 10] for the details.

An inductive procedure described in [**21**, Chapter 10] and [**324**, Chapter 13] explains how to go from here to the full numerical irreducible decomposition of V. These references also explain how one get information about the multiplicities of the components.

The Bertini Theorems. The methods described above work because there is a well-developed theory of what happens when we intersect a variety V with a generic affine subspace L. The basic results are due to Eugenio Bertini (1846-1933), a student of Cremona. His theorems, originally stated for $\mathbb{P}^n$ in 1882, used

the language of linear systems. Kleiman's article [**236**] gives a nice overview of what Bertini did and what his results mean in modern language.

To give a quick overview of the Bertini theorems, suppose that $V \subseteq \mathbb{C}^n$ has dimension ≥ 1 and let $H \subseteq \mathbb{C}^n$ be an affine hyperplane. If V is irreducible and H is generic, then the Bertini theorems tell us that

- $V \cap H$ is smooth outside $V_{\text{sing}} \cap H$.
- If $\dim(V) \geq 2$, then $V \cap H$ is irreducible.

Also, for arbitrary V and generic H, we know that

- $\dim(V \cap H) = \dim(V) - 1$.
- $\deg(V \cap H) = \deg(V)$.

It follow that if a variety V is irreducible of dimension $k > 1$, then for generic affine hypersurfaces $H_1, \ldots, H_k$,

$$V \cap H_1 \cap \cdots \cap H_k \text{ consists of } \deg(V) \text{ smooth points.}$$

Since $L = H_1 \cap \cdots \cap H_k$ is generic of codimension k, we see that the structure of the witness point set $V \cap L$ is a consequence of the Bertini theorems. Intersection with linear spaces is a key part of Bertini. We now see where the program got its name!

Four Examples. Rather than describe further features of software for numerical algebraic geometry, we instead give four extended examples to illustrate how these software packages are used in practice.

Example 2.14 (Lagrange Multipliers). Recall that optimizing $x^3 + cxyz - z^2$ on the sphere $x^2 + y^2 + z^2 = 1$ in $\mathbb{R}^3$ leads to the system of equations:

$$\begin{aligned} 3x^2 + cyz &= \lambda \cdot 2x \\ cxz &= \lambda \cdot 2y \\ cxy - 2z &= \lambda \cdot 2z \\ x^2 + y^2 + z^2 &= 1. \end{aligned} \tag{2.17}$$

For $c = 2$, one solution is $(x, y, z, \lambda) = (-\frac{2}{3}, \frac{1}{3}, \frac{2}{3}, -\frac{4}{3})$. However, a program such as Bertini would write this solution using floating point complex numbers:

```
-6.666666666666664e-01 -6.661338147750939e-16
3.333333333333339e-01 -6.383782391594650e-16
6.666666666666674e-01 -1.276756478318930e-15
-1.333333333333333e+00 -2.220446049250313e-16
```

real part | imaginary part

The real parts are reasonable, while the imaginary parts look odd until we realize that exponents such as "`e-16`" are telling us that the imaginary parts are zero!

There are several ways to solve (2.17) using Bertini. When $c = 2$, the simplest way is to use the start system $x^2 = 1, y^2 = 1, z^2 = 1, \lambda^2 = 1$ with 16 solutions. Bertini tracks $2^4 = 16$ paths, which behave as follows:

$$\begin{array}{l} \text{8 paths go to 8 nonsingular solutions} \\ \text{4 paths go to 2 solutions, each of multiplicity 2} \\ \text{4 paths go to } \infty. \end{array} \tag{2.18}$$

This gives 10 solutions, as reported in Example 2.2.

For such a simple system, what we did in (2.18) is all that is needed. But by looking more deeply, we can get insight into how BERTINI can solve more complicated systems. Recall from Example 2.2 that the system (2.17) has 12 rather than 10 solutions when $c = 2.1$ and $c = 2.01$. So $c = 2$ has *special features* that differ from the *generic behavior* of the system. Furthermore, we expect the generic case to be easier to solve since there may be fewer multiple solutions or fewer paths that go to ∞, both of which can be hard to deal with numerically.

In BERTINI, this is implemented in a two step strategy:

- Step1: Solve the system (2.17) with a random choice of the parameter $c \in \mathbb{C}$. We use the same start system as in (2.18), so there are still 16 paths to follow. But the paths behave differently:

  10 paths go to 10 nonsingular solutions
  4 paths go to 2 solutions, each of multiplicity 2
  2 paths go to ∞.

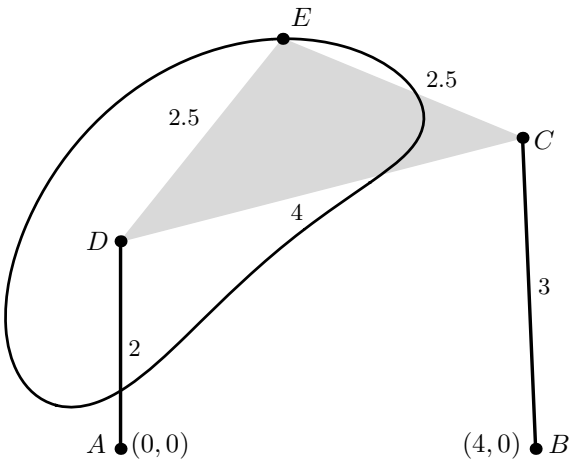
- Step 2: Solve (2.17) with $c = 2$, but now the start system is (2.17) with the c used in Step 1, and the solutions of the start system are the 14 finite solutions found in Step 1. When we follows these 14 paths, we obtain:

  8 paths go to 8 nonsingular solutions
  4 paths go to 2 solutions, each of multiplicity 2
  2 paths go to ∞.

  We get the same 10 solutions as in Example 2.18 with fewer paths to follow.

Step 2 is what BERTINI calls a *parameter homotopy* (see [**21**, Chapter 6] for details). Going from 16 paths to 14 paths is not very impressive, but the reduction can be much more dramatic, as happens in next example. ◁▷

EXAMPLE 2.15 (The Four-Bar Mechanism). In $\mathbb{R}^2$, consider a triangle CDE with edge lengths shown below and line segments AD, BC with indicated lengths. When we fix the points A, B and allow the triangle to move, the point E traces out a nice curve:



This is a classic *four-bar mechanism*, with bars AD, CD, BC and AB (not drawn) and a rigid body (the triangle CDE) attached to one of the bars.

Our first task is to find the equation of the curve traced out by E. If we set $C = (u, v), D = (c, s), E = (x, y)$, then the lengths in the above picture give five equations:

$$\begin{aligned} c^2 + s^2 &= 2^2 \\ (u-4)^2 + v^2 &= 3^2 \\ (u-c)^2 + (v-s)^2 &= 4^2 \\ (x-c)^2 + (y-s)^2 &= (5/2)^2 \\ (x-u)^2 + (y-v)^2 &= (5/2)^2. \end{aligned} \tag{2.19}$$

These equations are not sufficient since they ignore the *orientation* of the triangle, which we want to preserve. The vectors

$$C - D = (u - c, v - s),\ E - D = (x - c, y - s)$$

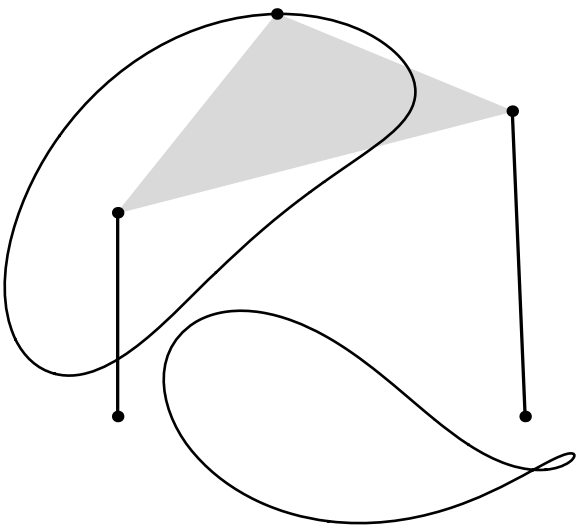
determine a parallelogram with signed area $+6$ when we go counterclockwise from $C - D$ to $E - D$. This gives the equation

$$\det \begin{pmatrix} u-c & v-s \\ x-c & y-s \end{pmatrix} = 6. \tag{2.20}$$

Combining this with (2.19) and eliminating u, v, c, s gives the sextic equation

$$\begin{aligned} &2025{-}2682x{-}1475x^2{-}2712x^3{+}2864x^4{-}768x^5{+}64x^6{-} \\ &5724y{+}8832xy{-}4320x^2y{+}1536x^3y{-}192x^4y{+}5949y^2{-} \\ &2712xy^2{+}3680x^2y^2{-}1536x^3y^2{+}192x^4y^2{-}4320y^3{+} \\ &1536xy^3{-}384x^2y^3{+}816y^4{-}768xy^4{+}192x^2y^4{-}192y^5{+}64y^6 = 0. \end{aligned} \tag{2.21}$$

The resulting picture has a surprise:

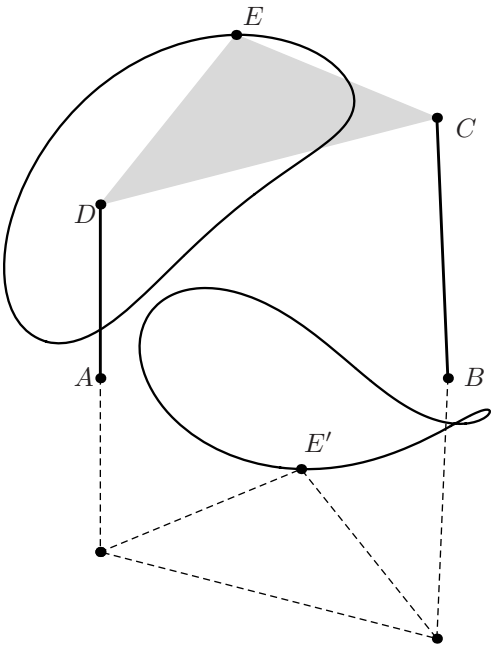


The real points of (2.21) consist of two connected components!

The Other Connected Component. To explain the other component, first observe that the equations (2.19) have two symmetries:

- Flip about the bottom edge of the triangle.
- Flip about the x-axis.

These symmetries are orientation-reversing. If we apply both to the gray triangle, we get the dashed triangle shown below. It has same orientation as the gray triangle and hence satisfies the signed area equation (2.20). This means that the point E' on the dashed triangle traces out the other component!



The Nine-Point Problem. A four-bar mechanism has nine degrees of freedom:

4: the two points A and B in $\mathbb{R}^2$
3: the three edge lengths of the triangle CDE
2: the two lengths of the line segments AD and BC.

It follows that if we specify nine points on the curve, there are now zero degrees of freedom. Intuitively, this means that there should be only finitely many four-bar mechanisms whose curve goes through the nine points. In 1923, Alt [**7**] posed the *nine-point problem* of finding a four-bar mechanism whose curve goes through nine arbitrarily chosen points in the plane.

In 1992, Wampler, Morgan and Sommese [**356**] used homotopy continuation to show that up to known symmetries, there are 1442 four-bar mechanisms whose curve goes through nine given points in $\mathbb{R}^2$. Computing this number was one of the first big successes of numerical algebraic geometry.

The nine-point problem is discussed in the books [**21**] and [**324**]. Following their approach, we change coordinates so that one of the nine points is the origin and replace coordinates x, y in the plane with *isotropic coordinates* $z = x+iy, z^* = x - iy$. Thus $z^* = \overline{z}$ precisely when x, y are real. Then the eight remaining points lead to a system of eight equations $f_1 = \cdots = f_8 = 0$ in eight unknowns $a_1, a_1^*, b_1, b_1^*, a_2, a_2^*, b_2, b_2^*$. Here, a_1, a_1^*, a_2, a_2^* are the isotropic coordinates of A, B, while b_1, b_1^*, b_2, b_2^* have more complicated descriptions [**21**, 5.5.2].

Four Bézout Numbers. When solving $f_1 = \cdots = f_8 = 0$ by homotopy continuation, there are four different ways to set up the start system:

- Each f_i is a polynomial of degree 7. When we work in $\mathbb{P}^8$, the standard start system of degree 7 polynomials would have $7^8 = 5{,}764{,}801$ solutions by Bézout's Theorem. This is a lot of paths to track!
- One virtue of isotropic coordinates is that the variables $a_1, a_1^*, b_1, b_1^*, a_2, a_2^*, b_2, b_2^*$ in f_i can be grouped in several ways. If we use the variable groups

$$\{a_1, a_1^*, b_1, b_1^*\} \text{ and } \{a_2, a_2^*, b_2, b_2^*\}, \tag{2.22}$$

  then f_i has degree 4 in each variable group. This leads to a start system on $\mathbb{P}^4 \times \mathbb{P}^4$ with 4,587,520 solutions by Bézout's Theorem for $(\mathbb{P}^4)^2$.
- For the variable groups $\{a_1, b_1\}, \{a_1^*, b_1^*\}, \{a_2, b_2\}, \{a_2^*, b_2^*\}$, f_i has degree 2 in each group. When we work in $\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^2$, this gives a start system with 645,120 solutions by Bézout's Theorem for $(\mathbb{P}^2)^4$.
- The polynomials $f_1, \ldots, f_8$ all have the same Newton polytope $P \subseteq \mathbb{R}^8$. By the `LATTICE_VOLUME` command in POLYMAKE [**9**], the normalized volume of P is 79,135. This leads to a start system with $\mathrm{vol}(P) = 79{,}135$ solutions by Bézout's Theorem for the toric variety X_P.

BERTINI can work in $\mathbb{P}^8$, $(\mathbb{P}^4)^2$ and $(\mathbb{P}^2)^4$ but not in X_P, though PHCPACK can.

We ran this example in BERTINI using the origin and eight randomly chosen points, with variable groups (2.22). This might seem inefficient (4,587,520 paths to track), but by using the BERTINI method called *regeneration*, the number of paths to follow can be dramatically reduced.

The rough idea of regeneration begins with a start system $(g_1, \ldots, g_8)$, where each g_i is a carefully chosen product of linear polynomials. But rather than go directly to the target system $(f_1, \ldots, f_8)$, we proceed in eight stages called *levels*. In Level 0, we homotop $(g_1, \ell_2 \ldots, \ell_8)$ to $(f_1, \ell_2 \ldots, \ell_8)$, where each ℓ_i is linear. In Level 1, we use the finite solutions of Level 0 and a homotopy for each linear factor of g_2 to find solutions of the start system $(f_1, g_2, \ell_3, \ldots, \ell_8)$. We then homotop this to $(f_1, f_2, \ell_3, \ldots, \ell_8)$. We continue until Level 7, where we finish with the solutions of $(f_1, \ldots, f_8)$. See [**21**, 5.4] for the details of how this works.

For the nine-point problem, regeneration gives the following BERTINI output:

```
|level| paths | nonsing endpts | total discarded | sing endpts | inf endpts | higher dim'l | other bad endpts
-------------------------------------------------------------------------------------------------------------
| 0   | 8     | 8              | 0               | 0           | 0          |  0           | 0
| 1   | 64    | 61             | 3               | 0           | 1          |  2           | 0
| 2   | 488   | 352            | 136             | 28          | 104        |  4           | 0
| 3   | 2816  | 1528           | 1288            | 346         | 936        |  6           | 0
| 4   | 11840 | 4688           | 7152            | 2100        | 5052       |  0           | 0
| 5   | 33064 | 9804           | 23260           | 7576        | 15684      |  0           | 0
| 6   | 55240 | 12176          | 43064           | 13304       | 29760      |  0           | 0
| 7   | 48703 | 8652           | 40051           | 11147       | 28904      |  0           | 0
-------------------------------------------------------------------------------------------------------------
|total| 152223
```

Only 152,223 paths were followed. The key feature of the output is the underlined number 8652. This is the number of the nonsingular solutions (possibly complex) of the nine-point equations. Because of a known six-fold symmetry for solutions of the nine-point problem, it follows that there are at most

$$8652/6 = 1442 \text{ distinct four-bar mechanisms through 9 points in } \mathbb{R}^2.$$

In 1992, the authors of [**356**] used a precursor of BERTINI written in Fortran on an IBM 3081. The computation took 332 hours of CPU time. In 2018, the same

computation takes about 60 hours using BERTINI on a laptop and under 2 hours using the parallel version of BERTINI on a cluster. ◁▷

EXAMPLE 2.16 (Parameter Estimation in a HIV Model). The 2016 paper *Numerical algebraic geometry for model selection and its application to the life sciences* by Gross, Davis, Ho, Bates and Harrington [**179**] discusses model validation, model selection, and parameter estimation using tools from numerical algebraic geometry.

We will focus on *parameter estimation* in a model that describes how the HIV virus infects $CD4^+$ T cells and macrophages, which are types of white blood cells in our immune system. Table 1 shows the model used in [**179**]. In what follows, we write "T cell" instead of "$CD4^+$ T cell" for simplicity.

Description	Interaction	Parameter value
Generation of new T cells	$\emptyset \xrightarrow{s_1} T$	10
Generation of new macrophages	$\emptyset \xrightarrow{s_2} M$	1.5×10^{-1}
Proliferation of T cells by presence of pathogen	$T + V \xrightarrow{k_1} (T + V) + T$	2×10^{-3}
Infection of T cells by HIV	$T + V \xrightarrow{k_2} T_i$	3×10^{-3}
Proliferation of M by presence of pathogen	$M + V \xrightarrow{k_3} (M + V) + M$	7.45×10^{-4}
Infection of M by HIV	$M + V \xrightarrow{k_4} M_i$	5.22×10^{-4}
Proliferation of HIV within T cell	$T_i \xrightarrow{k_5} V + T_i$	5.37×10^{-1}
Proliferation of HIV within macrophage	$M_i \xrightarrow{k_6} V + M_i$	2.85×10^{-1}
Natural death of T cells	$T \xrightarrow{\delta_1} \emptyset$	0.01
Natural death of infected T cells	$T_i \xrightarrow{\delta_2} \emptyset$	0.44
Natural death of macrophages	$M \xrightarrow{\delta_3} \emptyset$	6.6×10^{-3}
Natural death of infected macrophages	$M_i \xrightarrow{\delta_4} \emptyset$	6.6×10^{-3}
Natural death of HIV	$V \xrightarrow{\delta_5} \emptyset$	?

TABLE 1. HIV Model

The interactions in Table 1 are easy to explain. For example, consider

$$\begin{aligned} T + V &\xrightarrow{k_1} (T + V) + T \\ T + V &\xrightarrow{k_2} T_i \\ T_i &\xrightarrow{k_5} V + T_i. \end{aligned}$$

The first line is the creation of more T cells when the virus is detected, the second line is infection of T cells, and the third line is when infected T cells release more copies of the virus. The story for macrophages is mostly similar. One difference is that the death rate δ_2 for infected T cells is much larger than the death rate δ_1 for uninfected ones, while for macrophages, the death rates δ_3 and δ_4 are the same since macrophages can be long-lived virus reservoirs. The final line of the table does not specify a value of the parameter δ_5. The goal is to estimate δ_5 using data.

The Mathematical Model. Table 1 is an example of a *biochemical reaction network*. Such networks will be studied in detail in Chapter 5. To model the

network in Table 1, we introduce variables

$$\begin{aligned} T &= \text{uninfected T cells} \\ T_i &= \text{infected T cells} \\ M &= \text{uninfected macrophages} \\ M_i &= \text{infected macrophages} \\ V &= \text{HIV virus population.} \end{aligned}$$

Using the Law of Mass Action (to be explained in Chapter 5), we get a system of ordinary differential equations:

$$\begin{aligned} \tfrac{d}{dt}T &= s_1 + k_1TV - k_2TV - \delta_1 T \\ \tfrac{d}{dt}T_i &= k_2TV - \delta_2 T_i \\ \tfrac{d}{dt}M &= s_2 + k_3MV - k_4MV - \delta_3 M \\ \tfrac{d}{dt}M_i &= k_4MV - \delta_4 M_i \\ \tfrac{d}{dt}V &= k_5T_i + k_6M_i - \delta_5 V. \end{aligned} \tag{2.23}$$

Steady States. Stable populations occur when

$$\tfrac{d}{dt}T = \tfrac{d}{dt}T_i = \tfrac{d}{dt}M = \tfrac{d}{dt}M_i = \tfrac{d}{dt}V = 0.$$

Over $\mathbb{C}$, the polynomials on the right-hand side of (2.23) define the *steady state variety*

$$V^{ss} \subseteq \mathbb{C}^5 \times \mathbb{C} = \mathbb{C}^6,$$

where the variables T, T_i, M, M_i, V live in $\mathbb{C}^5$ and the unknown parameter δ_5 lives in the final $\mathbb{C}$. In [**179**], MACAULAY2 is used to decompose V^{ss} into irreducible components

$$V^{ss} = V_m \cup V_e,$$

with *main component* $V_m = \mathbf{V}(f_1, f_2, f_3, f_4, f_5)$, where

$$\begin{aligned} f_1 &= 5742\,M - 2453\,M_i - 130500 \\ f_2 &= 259908\,T_i - 46607\,M_i + 4840000\,\delta_5 - 20200500 \\ f_3 &= 17721\,T + 46607\,M_i - 4840000\,\delta_5 + 2479500 \\ f_4 &= 484000\,V\delta_5 - 184547\,M_i + 4840000\,\delta_5 - 20200500 \\ f_5 &= 2453\,M_iV - 72600\,M_i + 130500\,V, \end{aligned} \tag{2.24}$$

and *extinction component*

$$V_e = \mathbf{V}(V, M_i, 11\,M - 250, T_i, T - 1000).$$

For the symbolic computation, the floating point numbers in Table 1 were converted into rational numbers.

We like the extinction component because the virus is gone ($V = 0$) and the T cells and macrophages are all healthy ($T_i = M_i = 0$). But medically, the main component is more interesting, since a positive solution of $f_1 = \cdots = f_5 = 0$ indicates a stable HIV infection that coexists with stable T cell and macrophage populations (infected and uninfected).

Estimating δ_5 From Data. Suppose we have data

$$(T, T_i, M, M_i, V) = (319.5, 46.85, 81.09, 136.67, 21.21).$$

We will use this vector, denoted *data*, to estimate the parameter δ_5. A naive approach would to substitute $(T, T_i, M, M_i, V) = data$ into (2.24) and solve for δ_5. This is indeed naive, since the substitution results in an inconsistent set of equations. Remember that we rarely have exact solutions in the real world!

It follows that the main component $V_m \subseteq \mathbb{C}^6$ is disjoint from the *data line* $L \subseteq \mathbb{C}^6$ given by

$$L = \{(data, u) \mid u \in \mathbb{C}\} = \{data\} \times \mathbb{C} \subseteq \mathbb{C}^6.$$

We find the best estimate for δ_5 by minimizing the weighted distance

$$d^2 = \tfrac{(T-319.5)^2}{\sigma_1^2} + \tfrac{(T_i-46.85)^2}{\sigma_2^2} + \tfrac{(M-81.09)^2}{\sigma_3^2} + \tfrac{(M_i-136.67)^2}{\sigma_4^2} + \tfrac{(V-21.21)^2}{\sigma_5^2} + \tfrac{(\delta_5-u)^2}{\sigma_6^2},$$

between a point $(T, T_i, M, M_i, V, \delta_5) \in V_m$ and a point $(data, u) \in L$. In this formula, the denominators σ_i^2 come from Gaussian distributions added to create a statistical model. Following [**179**], we assume $\sigma_i^2 = 2$ for all i.

Finding the Global Minimum. Many standard numerical methods will find a local minimum of d^2. Using numerical algebraic geometry, we will find the point $(T, T_i, M, M_i, V, \delta_5, u) \in V_m \times \mathbb{C}$ at which d^2 attains its global minimum. The value of δ_5 at this point gives the desired estimate.

Since V_m is defined by $f_1 = \cdots = f_5 = 0$, this is constrained minimization. By Lagrange multipliers, the minimum of d^2 on $\mathbb{C}^6 \times \mathbb{C}$ with $f_1 = \cdots = f_5 = 0$ occurs at a critical point of $d^2 + \sum_{i=1}^5 \lambda_i f_i$. Solving this in BERTINI gives 16 solutions, three of which are real. Evaluating d^2 at the real solutions gives the estimate $\delta_5 = 2.99876$, which as noted in [**179**] is close to the known value $\delta_5 = 3$. ◁▷

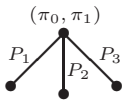
EXAMPLE 2.17 (Maximum Likelihood Estimation in Phylogenetics). Our final example is from the 2017 paper *Geometry of symmetric group-based models* by Korta and Kubjas [**237**]. In the introduction, the authors write

> The maximum likelihood estimation aims to find parameters that maximize the likelihood of observing the data for the given phylogenetic tree and phylogenetic model. In practice, numerical methods are used to find the MLE. However, as MLE is a non-convex optimization problem, there is no guarantee of outputting the global optimum. Since phylogenetic models are not necessarily compact, the MLE might even not exist.

One of the examples in [**237**] is a phylogenetic model due to Cavender, Farris and Neyman based on the tripod $K_{1,3}$. The paper uses numerical algebraic geometry to show that in this model, there are data vectors for which the MLE does not exist.

The Symmetric Group-Based Model. We have a root and three leaves, with probabilities π_0, π_1 at the root and transition matrices P_1, P_2, P_3 along the edges:

(2.25)



$$P_i = \begin{bmatrix} \frac{1}{2}+\frac{1}{2}e^{-2t_i} & \frac{1}{2}-\frac{1}{2}e^{-2t_i} \\ \frac{1}{2}-\frac{1}{2}e^{-2t_i} & \frac{1}{2}+\frac{1}{2}e^{-2t_i} \end{bmatrix},\ t_i \geq 0 \qquad \begin{matrix} \text{Group:} \\ \mathbb{Z}/2\mathbb{Z} = \{0, 1\} \end{matrix}$$

In this model, the root represents an extinct species with an unknown distribution (π_0, π_1) of some binary character. For example, DNA is built from four bases A, T, G and C. We get a binary character by looking at the distribution π_0 of *purines* (involved in A and G) and π_1 of *pyrimidines* (involved in C and T) in the original species. This species has evolved into three species represented by the leaves, and the transition matrices P_1, P_2, P_3 (also unknown) represent how the

binary character is transmitted to each new species. The parameters t_1, t_2, t_3 record how much time has passed.

We only have access to the three living species. Checking the binary character for each of these leads to eight possibilities. For $i, j, k \in \{0, 1\}$, let p_{ijk} denote the probability of observing character i in the first species, j in the second, and k in the third. The model tells us that

$$p_{ijk} = \pi_0 (P_1)_{0i} (P_2)_{0j} (P_3)_{0k} + \pi_1 (P_1)_{1i} (P_2)_{1j} (P_3)_{1k},$$

where rows and columns of P_1, P_2, P_3 are indexed by $0, 1$. The p_{ijk} satisfy the obvious restrictions $p_{ijk} \geq 0$ and $p_{000} + \cdots + p_{111} = 1$.

In [**237**], the authors show that the p_{ijk} that occur are characterized by equations and inequalities as follows:

$$\begin{array}{l} 1 \text{ linear equation } p_{000} + \cdots + p_{111} = 1 \\ 3 \text{ quadratic equations} \\ 8 \text{ linear weak inequalities } p_{ijk} \geq 0 \\ 7 \text{ strict linear inequalities} \\ 4 \text{ quadratic weak inequalities.} \end{array} \tag{2.26}$$

The full set of equations and inequalities are written out in [**237**, Example 2].

The Maximum Likelihood Estimate. Given a data point $(u_{000}, \ldots, u_{111})$, the *ML estimate* is the point $(p_{000}, \ldots, p_{111})$ of the model that "best explains" the data. Here, "best" means the point that gives the *global maximum* of the function

$$\prod_{i,j,k \in \{0,1\}} p_{ijk}^{u_{ijk}} \tag{2.27}$$

on the model. Standard numerical methods in statistics give a *local maximum* that is not guaranteed to be the global maximum. This is where numerical algebraic geometry enters the picture.

The Maximum Likelihood Degree. For the model (2.25), we get a complex algebraic variety by taking its Zariski closure. This means using the equations in (2.26) and ignoring the inequalities. The number of complex critical points of the likelihood function (2.27) on the Zariski closure is the *ML degree* of the model, which equals 92 in this case.

However, the weak inequalities in (2.26) mean that the model has a boundary. Since a global maximum can occur at a boundary point, we need to take these into account. The boundary may have several components, depending on which inequalities become equalities. So there are multiple Zariski closures to consider, each with its own ML degree.

An Example With No MLE. Consider the data

$$(u_{000}, \ldots, u_{111}) = (100, 11, 85, 55, 56, 7, 75, 8).$$

For these values of u_{ijk}, the MLE estimate is maximal value of (2.27) on the model (2.26). The analysis in [**237**, Example 8] involves 44 varieties, one coming from the equations in (2.26) and the other 43 coming from boundary components where various inequalities become equalities.

For each of these 44, the authors set up a Lagrange multipliers problem that they solved using PHCpack. Of the 167 critical points that were found, 49 were real and positive. Plugging into (2.27) gives candidates for the ML estimate.

For a boundary component where some inequalities become equalities, the other inequalities are ignored in computing the Zariski closure. Hence MLE candidates on such components need to be checked against the remaining inequalities. For the candidates coming from the seven largest values of (2.27), here is what happens:

- In four cases, some of the inequalities in (2.26) become negative. It follows that these candidates lie outside the topological closure of the model and hence have no meaning.
- In two more cases, weak inequalities derived from (2.26) become negative. So these candidates also lie outside the topological closure of the model.

This leaves the seventh top candidate, which also fails to satisfy (2.26). But unlike the previous six cases, it is contained in the topological closure of the model. The point corresponds to root distribution $(\pi_0, \pi_1) = (1, 0)$ and time values $(t_1, t_2, t_3) = (.665, +\infty, .262)$. Thus

$$\underbrace{\text{point for } (.665, +\infty, .262)}_{\text{not in the model}} = \lim_{t_2 \to +\infty} \underbrace{\text{point for } (.665, t_2, .262)}_{\text{in the model}}.$$

It follows that the MLE does not exist as a point of the model.

As the authors of [**237**] emphasize, local methods would never be able to reach this conclusion—one needs a global approach based on numerical algebraic geometry to discover the subtle behavior of the model (2.25). ◁▷

Final Comments. The goal of this section was to introduce homotopy continuation, with many omissions, and give four examples of numerical algebraic geometry in action. But there is much more that can be done with these methods, as you will see in the next section.

2.3. Applications of Sampling in Numerical Algebraic Geometry

by Jonathan Hauenstein

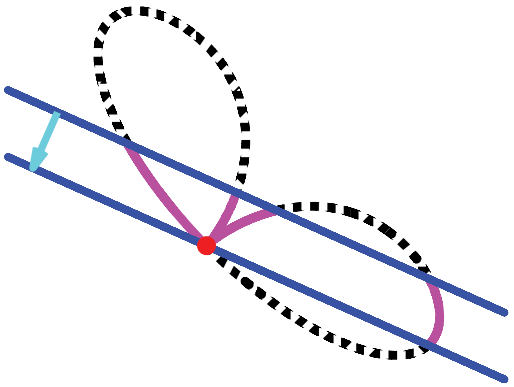
Many algorithms in numerical algebraic geometry start with a witness point set $V \cap L$ for an irreducible variety $V \subseteq \mathbb{C}^n$ of dimension k and degree d, and utilize homotopy continuation to deform the affine subspace L. This idea has already been implicitly utilized in various computations in Section 2.2 such as computing monodromy groups, deforming parameters in Example 2.14, and using regeneration for solving in Example 2.15. This section will explore other numerical algebraic geometric algorithms to obtain additional information about V.

Membership Testing and Local Multiplicity. When $V = \mathbf{V}(G)$ for some polynomial system G, a given point x^* is contained in V if and only if $G(x^*) = 0$. Witness point sets provide an alternative approach for deciding membership in V without computing such a polynomial system G which also yields the local multiplicity [**325**]. Recall that L is a general affine subspace of codimension k and let L_{x^*} be a general affine subspace of codimension k passing through x^*. Using homotopy continuation, one simply homotops $V \cap L$ to $V \cap L_{x^*}$. Then, $x^* \in V$ if and only if $x^* \in V \cap L_{x^*}$. The local multiplicity of x^* with respect to V is equal to the number of homotopy paths which end at x^* and is positive if and only if $x^* \in V$.

EXAMPLE 2.18. Let $V \subseteq \mathbb{C}^2$ be the quartic curve defined by

$$x_1^4 + x_2^4 - (x_1 + x_2)(x_1 - x_2)^2 = 0$$

and $L \subseteq \mathbb{C}^2$ be the affine linear subspace defined by $4x_1 + 9x_2 = 3$. The witness point set $V \cap L$ consists of 4 points. In order to compute the local multiplicity at $x^* = (0, 0)$, consider the linear slice L_{x^*} defined by $4x_1 + 9x_2 = 0$. The homotopy deforming $V \cap L$ to $V \cap L_{x^*}$ yields one path ending at a point which is not x^*, namely $(7605/6817, -3380/6817) \approx (1.1156, -0.4958)$, with the other three ending at x^* showing that the local multiplicity of x^* is 3. The following picture shows the 4 paths tracked along the quartic curve as L is deformed to L_{x^*} along V:



◁▷

Witness Point Sets of Projections. The geometric counterpart to elimination theory described in Chapter 1 is to compute images of varieties under projections. For the projection map $\pi(x) = (x_1, \ldots, x_\ell)$, Chevalley's Theorem provides that $\pi(V) \subseteq \mathbb{C}^\ell$ is a constructible set so that the closure under the Euclidean and Zariski topologies are equal, i.e., $\overline{\pi(V)}$ is a variety. In particular, if $m = \dim \overline{\pi(V)}$ and M is a random affine subspace of $\mathbb{C}^\ell$ of codimension m, we aim to compute a witness point set $\overline{\pi(V)} \cap M = \pi(V) \cap M$ from $V \cap L$ using the approach of [**195**].

Since $k = \dim V \geq \dim \overline{\pi(V)} = m$, we can select L_π to be a random affine subpace of codimension k inside of the codimension m affine space $\pi^{-1}(M) \subseteq \mathbb{C}^n$. Using homotopy continuation, one simply homotops $V \cap L$ to $V \cap L_\pi$. Then,

$$\pi(V \cap L_\pi) = \pi(V) \cap M$$

is a witness point set for $\overline{\pi(V)}$ thereby yielding the degree of $\overline{\pi(V)}$. Moreover, the generic fibers of $V \mapsto \pi(V)$ have dimension $k - m$ and degree $|V \cap L_\pi|/|\pi(V \cap L_\pi)|$.

EXAMPLE 2.19. Consider the curve $C = \overline{\pi(V)} \subseteq \mathbb{C}^2$ where $V \subseteq \mathbb{C}^3$ is the cubic curve defined by

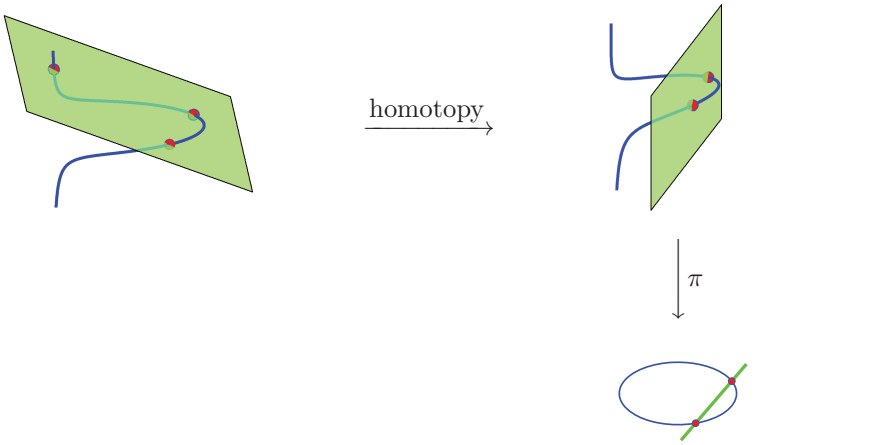
$$\begin{bmatrix} x(1+t^2) - (1-t^2) \\ y(1+t^2) - 2t \end{bmatrix} = 0 \quad \text{and} \quad \pi(x, y, t) = (x, y).$$

Since $\deg V = 3$, suppose that we start with a witness point set $V \cap L$ consisting of 3 points where $L \subseteq \mathbb{C}^3$ is plane defined by $4x - 2y + t = 3$. We aim to compute the witness point set $C \cap M$ where $M \subseteq \mathbb{C}^2$ is the line defined by $4x - 2y = 3$. Since

$\dim C = \dim V = 1$, we simply have $L_\pi = \pi^{-1}(M) = M \times \mathbb{C}$. The homotopy from $V \cap L$ to $V \cap L_\pi$ tracks 3 paths of which only 2 converge in $\mathbb{C}^3$ to the points

$$(x, y, t) = \left(\frac{6 \pm \sqrt{11}}{10}, -\frac{3 \mp 2\sqrt{11}}{10}, -\frac{2 \mp \sqrt{11}}{7} \right). \tag{2.28}$$

Since the images of these points under π are distinct, this shows that $\deg C = 2$ and $V \mapsto C$ is generically a one-to-one map. The following pictorially summarizes this computation:



◁▷

One can extend the membership test described above to decide membership in $\overline{\pi(V)}$ [**193**]. The following describes an application from [**197**] relating to 2×2 matrix multiplication. The book [**251**] provides extensive details relating matrix multiplication, tensors, and complexity.

EXAMPLE 2.20. Multiplication of 2×2 matrices is a bilinear map $\mathbb{C}^4 \times \mathbb{C}^4 \to \mathbb{C}^4$ which can be represented as the following tensor in $\mathbb{C}^4 \otimes \mathbb{C}^4 \otimes \mathbb{C}^4$:

$$\begin{aligned} T = {} & a_{11} \otimes b_{11} \otimes c_{11} + a_{12} \otimes b_{21} \otimes c_{11} + a_{11} \otimes b_{12} \otimes c_{12} + a_{12} \otimes b_{22} \otimes c_{12} + \\ & a_{21} \otimes b_{11} \otimes c_{21} + a_{22} \otimes b_{21} \otimes c_{21} + a_{21} \otimes b_{12} \otimes c_{22} + a_{22} \otimes b_{22} \otimes c_{22}. \end{aligned}$$

The 8 terms listed in T correspond with the 8 scalar multiplications needed to multiply to 2×2 matrices using the standard definition of matrix multiplication.

For $m \geq 1$, consider the variety

$$\sigma_m = \overline{\left\{ \sum_{i=1}^{m} \ell_1(a) \otimes \ell_2(b) \otimes \ell_3(c) \;\middle|\; \ell_i \text{ linear form on } \mathbb{C}^4 \right\}}.$$

The minimum m such that $T \in \sigma_m$ is called the *border rank* of T. The definition of T above shows that the border rank is at most 8. Strassen [**331**] showed that the border rank of T is at most 7 via the decomposition

$$\begin{aligned} T = {} & (a_{11} + a_{22}) \otimes (b_{11} + b_{22}) \otimes (c_{11} + c_{22}) + (a_{21} + a_{22}) \otimes b_{11} \otimes (c_{21} - c_{22}) \\ & + a_{11} \otimes (b_{12} - b_{22}) \otimes (c_{12} + c_{22}) + a_{22} \otimes (b_{21} - b_{11}) \otimes (c_{11} + c_{21}) \\ & + (a_{11} + a_{12}) \otimes b_{22} \otimes (c_{12} - c_{11}) + (a_{21} - a_{11}) \otimes (b_{11} + b_{12}) \otimes c_{22} \\ & + (a_{12} - a_{22}) \otimes (b_{21} + b_{22}) \otimes c_{11}. \end{aligned}$$

Thus, the border rank of T is 7 if and only if $T \notin \sigma_6$. By viewing σ_6 as the closure of an image under a projection map, numerical algebraic geometry was used in [**197**] to show $\deg \sigma_6 = 15{,}456$ by computing a witness point set for σ_6. Then, membership testing using this witness point set showed that $T \notin \sigma_6$. ◁▷

Two questions immediately arise from this last example. First, how can one verify the degree was computed correctly using numerical algebraic geometry? Second, can one use witness point sets to determine polynomials which vanish on a variety? These questions are addressed over the next four topics.

Trace Test. Given a subset W of a witness point set $V \cap L$ where V is irreducible, one aims to decide if $W = V \cap L$. One approach for deciding equality is via the *trace test* first described in [**326**]. Let M_t be a family of parallel linear spaces linearly parameterized by t with $M_0 = L$. The idea is to consider the sum of the solution paths starting at the points in W defined by $V \cap M_t$. One has a complete witness point set if and only if this sum of the solution paths is a linear function of t.

EXAMPLE 2.21. To demonstrate the essence of the trace test, consider the irreducible bivariate polynomial

$$f(x, y) = x^2 - xy + 2y^2 + 3x - 2y + 1$$

so that $f = 0$ defines an irreducible curve $C \subseteq \mathbb{C}^2$ of degree 2. The intersection of C with the parallel family of lines M_t defined by $y = 3x + t$ is described by

$$y = 3x + t \quad \text{and} \quad f(x, 3x + t) = 16x^2 + (11t - 3)x + (2t^2 - 2t + 1) = 0 \tag{2.29}$$

yielding two solution paths

$$(x_\pm(t), y_\pm(t)) = \frac{1}{32}\left(3 - 11t \pm \sqrt{-7t^2 + 62t - 55}, 9 - t \pm 3\sqrt{-7t^2 + 62t - 55}\right).$$

Clearly, both solution paths $(x_+(t), y_+(t))$ and $(x_-(t), y_-(t))$ are nonlinear (and not polynomial) in t showing that $\deg C > 1$. However, their sum

$$(x_+(t) + x_-(t), y_+(t) + y_-(t)) = \frac{1}{16}(3 - 11t, 9 - t)$$

is linear in t showing that $\deg C = 2$.

The value of $x_+(t) + x_-(t)$ is derived directly from f in (2.29) by simply observing that this sum is equal to the negative of the coefficient of x (linear in t) divided by the coefficient of x^2 (constant) by writing

$$f(x, 3x+t) = 16(x - x_+(t))(x - x_-(t)) = 16(x^2 - (x_+(t) + x_-(t)) \cdot x + x_+(t) \cdot x_-(t)).$$

This sum is called the *linear trace* which arises from the linear elementary symmetric function $e_1(a, b) = a + b$. One can also provide a similar statement using any nonconstant elementary symmetric function. For example, $e_2(a, b) = a \cdot b$ yields

$$(x_+(t) \cdot x_-(t), y_+(t) \cdot y_-(t)) = \frac{1}{16}(2t^2 - 2t + 1, t^2 - 9t + 9)$$

which is quadratic in t and again shows that $\deg C = 2$. Similar with $x_+(t) + x_-(t)$, the quadraticity of $x_+(t) \cdot x_-(t)$ is obtained from f in (2.29) where the value is equal to the coefficient of 1 (quadratic in t) divided by the coefficient of x^2 (constant).◁▷

From a numerical computational standpoint, one decides linearity of an analytic function by evaluating it at three (or more) points and testing linearity of the points.

EXAMPLE 2.22. The following illustrates using numerical algebraic geometry to compute the degree of the hypersurface $V \subseteq \mathbb{C}^{10}$ which is the closure of the bivariate cubic polynomials that can be written as a sum of three cubes of linears:

$$V = \overline{\left\{ a \in \mathbb{C}^{10} \;\middle|\; a_0 + a_1 x + a_2 y + \cdots + a_9 y^3 = \sum_{i=1}^{3} (b_{i0} + b_{i1} x + b_{i2} y)^3 \right\}}.$$

In other words, $V = \overline{\phi(\mathbb{C}^9)}$ where

$$\phi(b_1, b_2, b_3) = \begin{bmatrix} b_{10}^3 + b_{20}^3 + b_{30}^3 \\ 3(b_{10}^2 b_{11} + b_{20}^2 b_{21} + b_{30}^2 b_{31}) \\ 3(b_{10}^2 b_{12} + b_{20}^2 b_{22} + b_{30}^2 b_{32}) \\ 3(b_{10} b_{11}^2 + b_{20} b_{21}^2 + b_{30} b_{31}^2) \\ 6(b_{10} b_{11} b_{12} + b_{20} b_{21} b_{22} + b_{30} b_{31} b_{32}) \\ 3(b_{10} b_{12}^2 + b_{20} b_{22}^2 + b_{30} b_{32}^2) \\ b_{11}^3 + b_{21}^3 + b_{31}^3 \\ 3(b_{11}^2 b_{12} + b_{21}^2 b_{22} + b_{31}^2 b_{32}) \\ 3(b_{11} b_{12}^2 + b_{21} b_{22}^2 + b_{31} b_{32}^2) \\ b_{12}^3 + b_{22}^3 + b_{32}^3 \end{bmatrix}.$$

We consider intersecting V with the parallel family of lines M_t parameterized by

$$a = s \begin{bmatrix} 1 \\ -2 \\ -2 \\ -3 \\ -2 \\ -3 \\ 13 \\ 11 \\ 13 \\ -2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ -1 \\ -2 \\ 1 \\ 1 \\ 1 \\ 2 \\ 3 \\ 1 \end{bmatrix} + \frac{t}{4} \begin{bmatrix} 0 \\ -1 \\ -2 \\ -3 \\ -1 \\ -2 \\ -2 \\ -3 \\ 2 \\ 3 \end{bmatrix}.$$

Therefore, if $s = a_0$ is linear in t, then all coordinates are linear in t. For generic t, homotopy continuation finds that $V \cap M_t$ consist of 4 points. The following table shows the values of $s = a_0$ (rounded to 4 decimal places) for the 4 paths, say $s_1(t), \ldots, s_4(t)$, for some values of t:

t	$s_1(t)$	$s_2(t)$	$s_3(t)$	$s_4(t)$	$\sum_{i=1}^{3} s_i(t)$	$\sum_{i=1}^{4} s_i(t)$
-1	-0.4157	-0.3498	0.3421	1.6696	-0.4234	1.2462
0	-0.5646	-0.1876	0.2573	1.5219	-0.4948	1.0271
1	-0.7269	-0.0025	0.1492	1.3883	-0.5803	0.8080

From this data, we test for linearity as follows. If a function $r(t)$ is linear, then

$$r(1) + r(-1) - 2r(0) = 0. \tag{2.30}$$

For $r(t) = \sum_{i=1}^{3} s_i(t)$, we have $r(1) + r(-1) - 2r(0) = -0.0141 \neq 0$ showing that $\sum_{i=1}^{3} s_i(t)$ is not linear in t. However, for $r(t) = \sum_{i=1}^{4} s_i(t)$, (2.30) holds. By the randomness of M_t, the points $t = -1, 0, 1$ were general so we can conclude that $\sum_{i=1}^{4} s_i(t)$ is linear in t. Therefore, the trace test concludes $\deg V = 4$. ◁▷

In practice, the coefficients of M_t are chosen with random complex numbers and the trace test is numerically performed using various precisions to provide a robust test of linearity.

Estimating Degrees. The trace test is used to determine if a complete witness point set has been found. If the purported witness point set is not complete, the trace test does not provide any indication of how many points are missing. A statistical estimation of the degree can be used for such an indication. The following describes an estimation approach based on subset sampling [**192**] with other estimation methods provided in [**18**, **292**].

Suppose that W is a known subset of a witness point set $V \cap L$ where the aim is to approximate $d = \deg V = \#V \cap L$. An estimation of d is obtained by computing a uniformly at random subset $Q \subseteq V \cap L$ of size $1 \leq \ell \leq d$ and comparing the sizes of W, Q, and $W \cap Q$. If $W \cap Q \neq \emptyset$, the maximum likelihood estimate for d is

$$\frac{\#W}{\frac{\#W \cap Q}{\#Q}} = \frac{\#W \cdot \#Q}{\#W \cap Q}.$$

The main denominator in the first expression is the ratio of repeats which is an estimation of the relative size of W to $V \cap L$. Hence, if the ratio of repeats is large, this suggests that W almost contains all of $V \cap L$. Conversely, if the ratio of repeats is small, then there are many points in $V \cap L$ not contained in W.

The main challenge to employ such a statistical estimation is the ability to actually generate a uniformly at random subset Q of $V \cap L$. In [**192**], a monodromy loop, which is discussed in Section 2.2, starting with some subset of W is used to generate the set Q.

Example 2.23 (Estimating the Nine-Point Problem). In Section 2.2, regeneration was used to solve the nine-point problem for four-bar linkages. This computation confirmed the result from [**356**] of 1442 distinct four-bar mechanisms. We start with one known solution. At each stage, a monodromy loop, which moves the nine points in a loop, is performed starting with all known solutions so that $\#W = \#Q$. The following summarizes the number of known solutions, number of repeats, and an estimate of the total number of solutions when the number of repeats is not zero at each stage in our experiment.

stage	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$\#W$	1	2	4	8	16	32	64	125	239	429	740	1108	1364	1436	1442
$\#W \cap Q$	0	0	0	0	0	0	3	11	49	118	372	852	1292	1430	1442
estimate	–	–	–	–	–	–	1365	1420	1166	1560	1472	1441	1440	1442	1442

This confirms the number 1442 computed in Example 2.15. ◁▷

Vanishing Polynomials. Witness point sets provide a geometric description of a variety. In some instances, algebraic information in the form of polynomials that vanish on the variety is also of interest. This can be accomplished using sampling and interpolation.

When searching for vanishing polynomials with integer coefficients, one approach described in [**22**] is to utilize an integer relation algorithm such as LLL [**256**] and PSLQ [**154**] with one general point. Given a vector $z \in \mathbb{C}^\ell$, integer relation algorithms seek a vector $a \in \mathbb{Z}^\ell$ such that $a \cdot z = 0$. By taking z to be a vector obtained from evaluating polynomials $p_1(x), \ldots, p_\ell(x)$ at a general point which can be computed to arbitrary accuracy via adaptive precision computations,

one is searching for coefficients $a \in \mathbb{Z}^\ell$ such that $a_1 p_1(x) + \cdots + a_\ell p_\ell(x)$ vanishes on the variety.

EXAMPLE 2.24. Suppose one aims to compute the polynomial vanishing on the curve $C = \overline{\pi(V)} \subseteq \mathbb{C}^2$ in Example 2.19. Since the points in (2.28) were defined by intersecting C with the line $4x - 2y = 3$ having integer coefficients, integer relation algorithms would be influenced by this line. Thus, we simply homotop one point to $C \cap M'$ where M' is defined by $4x - 2y = 2\sqrt{2}$ to obtain the point

$$(x, y) = \frac{1}{5}(2\sqrt{2} + \sqrt{3}, 2\sqrt{3} - \sqrt{2}) \in C. \tag{2.31}$$

Even though we know that no linear polynomials vanish on C since $C \subseteq \mathbb{C}^2$ is an irreducible curve of degree 2, we use this as an illustrative case when searching for a nonexistent relation. By taking $z = (1, x, y)$, the PSLQ algorithm in MAPLE returns the integer relation vectors

$$a = (-102543, 80992, 69932) \quad \text{and} \quad a = (-6837579893, 2055056255, 12105959021)$$

using 16 and 32 digits, respectively. If we increase the number of digits to 200, one obtains

$$\begin{aligned} a = (&316264594137797502415297326205736048706654661115877081341496871 42, \\ &-113288418162609754421828242615578226870131457569293768604938682 4651, \\ &244323635446873166876043322328617827772845029688772602357899482 9436) \end{aligned}$$

which indicates that searching for a linear integer relation is futile.

Next, we search for the quadratic polynomial which defines C using the vector $z = (1, x, y, x^2, xy, y^2)$. When using 16 digits, one obtains

$$a = (-1, 0, 0, 1, 0, 1) \qquad \text{corresponding to} \qquad x^2 + y^2 - 1 = 0.$$

In this case, one can easily verify that this is the correct relation by substitution since C is parameterized by

$$x = \frac{1 - t^2}{1 + t^2} \qquad \text{and} \qquad y = \frac{2t}{1 + t^2}. \qquad \lhd\rhd$$

An alternative to using one point to search for integer relations is to utilize many points via classical interpolation with numerical linear algebra. As above, suppose that $p_1(x), \ldots, p_\ell(x)$ are given polynomials and one aims to obtain all $a \in \mathbb{C}^\ell$ such $a_1 p_1(x) + \cdots + a_\ell p_\ell(x)$ vanishes on a given variety V. This is equivalent to computing the null space of the matrix

$$\left[\begin{array}{cccc} p_1(x) & \cdots & p_\ell(x), & x \in V \end{array} \right]$$

Since the number of columns is ℓ, one only needs at most ℓ general points of V to determine the proper null space. In fact, one has an easy to check stopping criterion since adding a row evaluated at a general point either reduces the dimension of the null space by one or one already has computed the correct null space.

EXAMPLE 2.25. We again aim to compute polynomials vanishing on the curve $C = \overline{\pi(V)} \subseteq \mathbb{C}^2$ in Example 2.19. First, linears vanishing on C correspond with $a \in \mathbb{C}^3$ such that $a_1 + a_2 x + a_3 y = 0$ on C. Using the three known points from

(2.28) and (2.31), we have the matrix

	1	x	y
$\frac{1}{10}(6+\sqrt{11}, 3-2\sqrt{11})$	1	$\frac{6+\sqrt{11}}{10}$	$\frac{3-2\sqrt{11}}{10}$
$\frac{1}{10}(6-\sqrt{11}, 3+2\sqrt{11})$	1	$\frac{6-\sqrt{11}}{10}$	$\frac{3+2\sqrt{11}}{10}$
$\frac{1}{5}(2\sqrt{2}+\sqrt{3}, 2\sqrt{3}-\sqrt{2})$	1	$\frac{2\sqrt{2}+\sqrt{3}}{5}$	$\frac{2\sqrt{3}-\sqrt{2}}{5}$

Since the resulting 3×3 matrix has full rank, we again confirm that no linear polynomials vanish on C.

To search for quadratic polynomials using the basis $1, x, y, x^2, xy, y^2$, we compute additional points on C by moving the intersecting linear space. For example,

	1	x	y	x^2	xy	y^2
$\frac{1}{10}(6+\sqrt{11}, 3-2\sqrt{11})$	1	$\frac{6+\sqrt{11}}{10}$	$\frac{3-2\sqrt{11}}{10}$	$\frac{47+12\sqrt{11}}{100}$	$\frac{-4-9\sqrt{11}}{100}$	$\frac{53-12\sqrt{11}}{100}$
$\frac{1}{10}(6-\sqrt{11}, 3+2\sqrt{11})$	1	$\frac{6-\sqrt{11}}{10}$	$\frac{3+2\sqrt{11}}{10}$	$\frac{47-12\sqrt{11}}{100}$	$\frac{-4+9\sqrt{11}}{100}$	$\frac{53+12\sqrt{11}}{100}$
$\frac{1}{5}(2\sqrt{2}+\sqrt{3}, 2\sqrt{3}-\sqrt{2})$	1	$\frac{2\sqrt{2}+\sqrt{3}}{5}$	$\frac{2\sqrt{3}-\sqrt{2}}{5}$	$\frac{11+4\sqrt{6}}{25}$	$\frac{2+3\sqrt{6}}{25}$	$\frac{14-4\sqrt{6}}{25}$
$\frac{1}{5}(2\sqrt{2}-\sqrt{3}, -2\sqrt{3}-\sqrt{2})$	1	$\frac{2\sqrt{2}-\sqrt{3}}{5}$	$\frac{-2\sqrt{3}-\sqrt{2}}{5}$	$\frac{11-4\sqrt{6}}{25}$	$\frac{2-3\sqrt{6}}{25}$	$\frac{14+4\sqrt{6}}{25}$
$(0, -1)$	1	0	-1	0	0	1
$\frac{1}{25}(24, -7)$	1	$\frac{24}{25}$	$\frac{-7}{25}$	$\frac{576}{625}$	$\frac{-168}{625}$	$\frac{49}{625}$

results in a 6×6 matrix of rank 5 with the null space spanned by the vector $a = (-1, 0, 0, 1, 0, 1)$ as expected. ◁▷

EXAMPLE 2.26. The following degree 4 homogeneous polynomial that defines V in Example 2.22 was obtained via sampling and interpolation:

$$\begin{aligned}
&144a_0a_3a_7a_9 - 48a_0a_3a_8^2 - 216a_0a_4a_6a_9 + 24a_0a_4a_7a_8 + 144a_0a_5a_6a_8 \\
&- 48a_0a_5a_7^2 - 48a_1^2a_7a_9 + 16a_1^2a_8^2 + 144a_1a_2a_6a_9 - 16a_1a_2a_7a_8 + 24a_1a_3a_4a_9 \\
&- 16a_1a_3a_5a_8 - 8a_1a_4^2a_8 + 24a_1a_4a_5a_7 - 48a_1a_5^2a_6 - 48a_2^2a_6a_8 + 16a_2^2a_7^2 \\
&- 48a_2a_3^2a_9 + 24a_2a_3a_4a_8 - 16a_2a_3a_5a_7 - 8a_2a_4^2a_7 + 24a_2a_4a_5a_6 + 16a_3^2a_5^2 \\
&- 8a_3a_4^2a_5 + a_4^4.
\end{aligned}$$

It is easy to prove that it vanishes on V via substitution since $V = \overline{\phi(\mathbb{C}^9)}$. ◁▷

Vanishing Polynomials on Witness Point Sets. A shortcoming of these approaches for recovering polynomials is the dimension of the vector space of polynomials under consideration. For example, the dimension of the vector space of polynomials in n variables of degree at most r is $\binom{n+r}{r}$. One way to lighten the workload is to consider the simpler problem of computing polynomials that vanish on a given witness point set which are restricted to the linear space used to construct it. In such a situation, one always obtains upper bounds in that every polynomial which vanishes on the variety restricts to a polynomial which vanishes on the given witness point set. If the variety is arithmetically Cohen-Macaulay, i.e., the Krull dimension of the coordinate ring is equal to its depth, then there is an equivalence between polynomials which vanish on the variety and vanish on a witness point set. In fact, one may turn this around and use it to test the arithmetically Cohen-Macaulayness of a variety [**117**].

EXAMPLE 2.27. Consider the twisted cubic curve $C = \overline{\phi(\mathbb{C})}$ and the quartic curve $D = \overline{\theta(\mathbb{C})}$ where

$$\phi(t) = (t, t^2, t^3) \quad \text{and} \quad \theta(t) = (t, t^3, t^4).$$

For $L \subseteq \mathbb{C}^3$ defined by $z = 3x - 3y + 1$, we aim to compare quadratic polynomials $f(x, y, z)$ vanishing on C and D with quadratic polynomials $g(x, y, 3x - 3y + 1)$ restricted to L vanishing on $C \cap L$ and $D \cap L$, respectively.

The set of quadratic polynomials vanishing on C is the three-dimensional vector space spanned by

$$f_1 = xy - z, \quad f_2 = x^2 - y, \quad f_3 = xz - y^2.$$

Restricting to L also yields a three-dimensional vector space spanned by

$$g_1 = xy - 3x + 3y - 1, \quad g_2 = x^2 - y, \quad g_3 = 3x^2 - 3xy - y^2 + x.$$

Since C is arithmetically Cohen-Macaulay, this space is exactly the space of quadratic polynomials restricted to L that vanish on the witness point set

$$C \cap L = \{(1, 1, 1), (-2 \pm \sqrt{3}, 7 \mp 4\sqrt{3}, -26 \pm 15\sqrt{3})\}.$$

In contrast, the set of quadratic polynomials vanishing on D is the one-dimensional vector space spanned by f_1 with corresponding g_1 when restricted to L. However, the space of quadratic polynomials restricted to L that vanish on

$$D \cap L = \left\{(\pm 1, \pm 1, 1), \frac{1}{2}(-3 \pm \sqrt{5}, -18 \pm 8\sqrt{5}, 47 \pm 21\sqrt{5})\right\}$$

is two-dimensional. It is spanned by g_1 and

$$160x^2 - 153xy - 21x + 20y^2 + 21y - 27.$$

Hence, D is not arithmetically Cohen-Macaulay since there is a two-dimensional set of quadratic polynomials restricted to L vanishing on $D \cap L$ while only a one-dimensional set of quadratic polynomials vanishing on D. ◁▷

Example 2.28. Returning to the border rank of 2×2 matrix multiplication from Ex. 2.20, one can first investigate polynomials vanishing on σ_6 by looking for polynomials which vanish on a witness point set $\sigma_6 \cap L$ and are restricted to L. For example, this shows that no nontrivial homogeneous polynomials of degree ≤ 18 vanish on $\sigma_6 \cap L$ and hence the same is true for σ_6. For comparison, the space of homogeneous polynomials of degree 18 in $\mathbb{C}^{64}$ has dimension

$$\binom{63 + 18}{18} = 456{,}703{,}981{,}505{,}085{,}600$$

while the corresponding computation on $\sigma_6 \cap L$ restricted to L only has dimension

$$\binom{4 + 18}{18} = 7{,}315.$$

In degree 19, there is a 64-dimensional space of polynomials vanishing on $\sigma_6 \cap L$. Representation theory was used in [**197**] to describe the corresponding polynomials vanishing on σ_6. To complete the story, an explicit polynomial vanishing on σ_6 was then constructed which did not vanish at the 2×2 matrix multiplication tensor T confirming that $T \notin \sigma_6$ thereby showing that the border rank of T is 7. ◁▷

Sums of Squares Modulo an Ideal. A polynomial $p \in \mathbb{R}[x_1, \ldots, x_n]$ is a sum of squares if there are $g_1, \ldots, g_\ell \in \mathbb{R}[x_1, \ldots, x_n]$ such that $p = g_1^2 + \cdots + g_\ell^2$. Clearly, if p is a sum of squares, then $p \geq 0$ on $\mathbb{R}^n$. In addition to $\mathbb{R}^n$, one can consider

certificates of nonnegative on other real varieties. For a variety $V \subseteq \mathbb{C}^n$, p is a sum of squares modulo the ideal $\mathbf{I}(V)$ if there are $g_1, \ldots, g_\ell \in \mathbb{R}[x_1, \ldots, x_n]$ such that

$$p = g_1^2 + \cdots + g_\ell^2 \mod \mathbf{I}(V)$$

which certifies $p \geq 0$ on $V \cap \mathbb{R}^n$.

If a generating set of $\mathbf{I}(V)$ is known, then sums of squares modulo $\mathbf{I}(V)$ where the degrees of the polynomials g_j are bounded can be determined using semidefinite programming (SDP). Rather than using defining polynomials of V, an alternative approach described in [**81**] is to utilize sample points. If $a_1, \ldots, a_m \in V$, a necessary condition for a polynomial p to be a sum of squares modulo $\mathbf{I}(V)$ is that there exists $g_1, \ldots, g_\ell \in \mathbb{R}[x_1, \ldots, x_n]$ such that $p(a_j) = g_1(a_j)^2 + \cdots + g_\ell(a_j)^2$ for $j = 1, \ldots, m$. This necessary condition where the degrees of g_j are bounded can also be solved via an SDP. The solution of this necessary condition can be checked to see if the purported decomposition is correct via testing at one extra general point in V. If the test fails, one adds additional sample points and solves again. A finite number of general points on V suffice to either compute a decomposition if it exists or show that no such decomposition exists when the degrees of g_j are bounded.

Example 2.29. To illustrate the advantage of using sample points, we utilize a modification of the trace ratio problem described in [**81**, §7.4]. Define

$$\mathcal{X}_{m,\ell} = \left\{ X \in \mathbb{R}^{m \times m} \mid X^T = X^2 = X, \operatorname{trace}(X) = \ell \right\}.$$

For symmetric matrices $A, B, C \in \mathbb{R}^{m \times m}$ where B is positive definite, compute

$$\gamma^\star = \max_{X \in \mathcal{X}_{m,\ell}} \frac{\operatorname{trace}(AX)}{\operatorname{trace}(BX)} + \operatorname{trace}(CX).$$

For $p(X, \gamma) = \operatorname{trace}(BX)(\gamma - \operatorname{trace}(CX)) - \operatorname{trace}(AX)$, an equivalent problem is

$$\gamma^\star = \min_{\gamma \in \mathbb{R}} \gamma \quad \text{such that} \quad p(\cdot, \gamma) \geq 0 \text{ on } \mathcal{X}_{m,\ell}$$

which is relaxed to

$$\gamma^\star_{\mathrm{SOS}} = \min_{\gamma \in \mathbb{R}} \gamma \quad \text{such that} \quad p(\cdot, \gamma) \text{ is a sum of squares mod } I(\mathcal{X}_{m,\ell}).$$

When searching for a sum of squares decomposition with $\deg g_j \leq 1$, the following from [**81**, Table 2] shows the improvement of using sampling over defining equations:

		Equations SDP			Sampling SDP		
m	ℓ	variables	constraints	time(s)	variables	constraints	time(s)
4	2	342	188	0.47	56	45	0.10
5	3	897	393	0.71	121	105	0.11
6	4	2062	738	1.34	232	210	0.15
7	5	4265	1277	3.62	407	378	0.19
8	6	8106	2073	9.06	667	630	0.34
9	7	14387	3198	23.83	1036	990	0.61
10	8	24142	4733	58.17	1541	1485	1.18

◁▷

Note that the sample points $a_1, \ldots, a_m$ used above were selected in V and not necessarily in $V \cap \mathbb{R}^n$. In other situations where required, the following describes an approach for sampling real points on a variety.

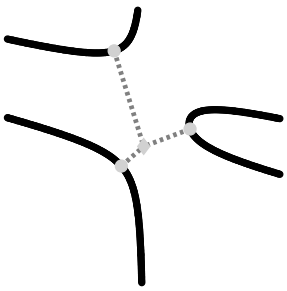
Sampling Real Points. In 1954, Abraham Seidenberg [**316**] stated what he called a "well-known" fact that if a variety contains real points, then there is a real

point on the variety of minimal Euclidean distance to the origin. This is an easy consequence of the extreme value theorem. Moreover, each connected component of the set of real points has such a minimal point thereby providing the basis for a method to compute a point on each connected component of the real points of a variety.

EXAMPLE 2.30. The cubic plane curve defined by

$$f(x,y) = 3x^3 + 24x^2y - 6x^2 + 48xy^2 - 24xy + 1 = 0$$

has three connected components where the minimizers of Euclidean distance to the origin on each connected component are shown in the following picture:



◁▷

One shortcoming of measuring distance to the origin is that the corresponding minimizers need not be unique. For example, every point on the unit circle $x^2 + y^2 = 1$ is equidistant to the origin. However, minimizing distance to a randomly selected point will yield a unique minimizer on each connected component. Thus, for a variety $V \subseteq \mathbb{C}^n$ and a randomly selected $p \in \mathbb{R}^n$, one aims to solve

$$\min\{\|x-p\|_2^2 \mid x \in V \cap \mathbb{R}^n\} \tag{2.32}$$

where $\|x-p\|_2^2 = (x_1-p_1)^2 + \cdots + (x_n-p_n)^2$. Similar to Section 2.2, one replaces solving (2.32) with computing the critical points of (2.32) via solving a polynomial system, namely computing $x \in V$ such that

$$\operatorname{rank}\begin{bmatrix} x-p & \operatorname{Jac}_V(x) \end{bmatrix} \leq n - \dim V \tag{2.33}$$

where $\operatorname{Jac}_V(x)$ is the Jacobian matrix of V at x.

EXAMPLE 2.31. The critical points for Example. 2.30 when $p = 0$ satisfy

$$f(x,y) = 0 \quad \text{and} \quad \operatorname{rank}\begin{bmatrix} x & \frac{\partial f}{\partial x} \\ y & \frac{\partial f}{\partial y} \end{bmatrix} \leq 1.$$

Enforcing the rank condition on the matrix by setting its determinant to zero yields

$$\begin{bmatrix} 3x^3 + 24x^2y - 6x^2 + 48xy^2 - 24xy + 1 \\ 24x^3 + 87x^2y - 24x^2 - 48xy^2 + 12xy - 48y^3 + 24y^2 \end{bmatrix} = 0.$$

This system has 7 solutions: three real minimizers plotted above in Example 2.30 which are approximately

$$(0.3408, 0.1277), \quad (-0.2164, 0.7026), \quad (-0.1641, -0.1441)$$

and four nonreal solutions which are approximately

$$(0.1971 \pm 0.2100\sqrt{-1}, 0.3842 \pm 0.0912\sqrt{-1}), (0.0483 \pm 0.1719\sqrt{-1}, -0.0336 \mp 0.2498\sqrt{-1}).$$ ◁▷

A drawback of (2.33) is that *every* singular point of V is considered a critical point. In [**299**], the rank condition (2.33) is enforced using determinants and an infinitesimal is used to perturb the system thereby yielding a finite subset of the critical points containing the global minimizer on each connected component. An alternative approach presented in [**191**] uses a null space, which is equivalent to the Fritz John conditions, to enforce (2.33) together with a homotopy in order to compute a finite subset of the critical points containing the global minimizer on each connected component. The following illustrates this homotopy-based approach.

EXAMPLE 2.32. Consider the hypersurface $V \subseteq \mathbb{C}^3$ defined by

$$f(x) = (x_2 - x_1^2)^2 + (x_1x_2 - x_3)^2 + (x_1x_3 - x_2^2)^2 = 0$$

for which $V \cap \mathbb{R}^3$ is the real part of the twisted cubic curve. Let $p = (-1, 2, 2)$. Every point in $V \cap \mathbb{R}^3$ is a singular point on V and hence a critical point of (2.32). For, say, $\epsilon = 1 + \sqrt{-1}$, we can compute a finite set of critical points via

$$H(x, \lambda; t) = \begin{bmatrix} f(x) - \epsilon t \\ \begin{bmatrix} x - p & \nabla f(x) \end{bmatrix} \lambda \end{bmatrix} = 0$$

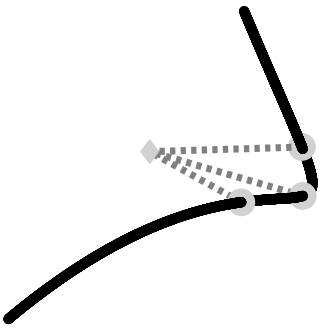
where $\lambda \in \mathbb{P}^1$. Solving at $t = 1$ yields 32 start points for the homotopy H. All 32 solution paths converge as $t \to 0$ yielding:

- (1 path each) 10 nonreal points which are smooth points on V
- (2 paths each) 3 real and 2 nonreal points which are singular points on V
- (3 paths each) 4 nonreal points which are singular points on V.

A complete accounting of the paths is $32 = 10 \cdot 1 + 5 \cdot 2 + 4 \cdot 3$. The 3 real points listed in increasing distance from p are approximately

$$(1.2138, 1.4734, 1.7884), \quad (-0.5969, 0.3563, -0.2127), \quad (0.2335, 0.0545, 0.0127)$$

which are the global minimum, a non-global local minimum, and a non-global local maximum, respectively, as observed in the following picture:



◁▷

It is easy to check that the selected choices of p and ϵ in Ex. 2.32 were generic in that one would obtain the same number of start points (32) and a similar accounting of the endpoints ($32 = 10 \cdot 1 + 5 \cdot 2 + 4 \cdot 3$) for generic choices of p and ϵ. The term *Euclidean distance degree* was coined in [**136**] to describe the generic number of critical points of the distance function for a variety. For example, the Euclidean distance degree of the smooth quartic hypersurface defined by $f(x) = \epsilon$ from Ex. 2.32 is 32.

By treating the point p as a parameter, one can utilize a parameter homotopy to efficiently solve many different instances of (2.32).

EXAMPLE 2.33. Consider sampling real points from $S \subseteq \mathbb{C}^3$ defined by

$$f(x,y,z) = 5(4x^2+9y^2+4z^2-4)^3 - 4z^3(80x^2+9y^2) = 0.$$

This surface is called *süss* which means "sweet" in German since the real part $S \cap \mathbb{R}^3$ resembles the boundary of a candy heart. For generic ϵ and p, the Euclidean distance degree of the smooth surface defined by $f = \epsilon$ is 96 which homotop to the following breakdown of critical points on S as $\epsilon \to 0$:

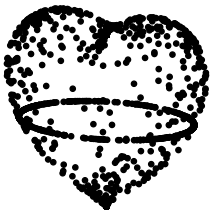
- (1 path each) 34 smooth points on S
- (3 paths each) 2 singular points $(0,0,\pm 1) \in S$
- (6 paths each) 4 points on the singular curve around the "waist" of S which is defined by

$$4x^2+9y^2-4 = z = 0 \tag{2.34}$$

- (8 paths each) 4 singular points $(\pm\sqrt{-1/19}, \pm\sqrt{80/171}, 0) \in S$

with $96 = 34 \cdot 1 + 2 \cdot 3 + 4 \cdot 6 + 4 \cdot 8$. The irreducible components of the singular set of S are the singular curve defined by (2.34) and the isolated points $(0,0,\pm 1)$. The other 4 singular points are embedded points on the singular curve.

The following picture plots the real critical points of the distance function to 200 randomly selected points $p \in [-2,2]^3$ obtained via parameter homotopies:



◁▷

Rather than use random choices of p which clusters sample points in high curvature regions, one can be more clever in the choice of test points to create a provably dense sampling of any compact subset of a real variety. This can be useful, for example, in topological data analysis [**135**].

Final Comments. The purpose of this section was to provide a sampling, with many omissions, of applications of computing and using sample points in numerical algebraic geometry. One common theme of these applications is the extraction of information about a variety from some particular set of points on the variety.

CHAPTER 3

Geometric Modeling

The interactions between geometric modeling, algebraic geometry, and commutative algebra are rich and varied. The first two sections of the chapter discuss the geometry and the algebra of geometric modeling. In the third section, Hal Schenck goes more deeply into the algebra.

3.1. Geometry of Modeling

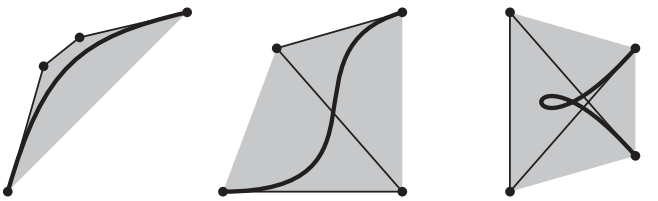
Geometric modeling involves curves in $\mathbb{R}^2$ and $\mathbb{R}^3$ and surfaces in $\mathbb{R}^3$. We will begin with curves and then move to the richer topic of surfaces.

Bézier Curves. As computer-aided design developed in the 1950s and 1960s, Paul de Casteljau of Citroën and Pierre Bézier of Renault independently created and implemented systems based on what we now call *Bézier curves*.[1] These curves use *blending functions* $\binom{n}{i}t^i(1-t)^{n-i}$ and *control points* $P_0, \dots, P_n \in \mathbb{R}^\ell$ to give the parametrization

$$\begin{aligned} &\Phi : [0,1] \longrightarrow \mathbb{R}^\ell \\ &\Phi(t) = \sum_{i=0}^{n} \binom{n}{i} t^i (1-t)^{n-i} P_i. \end{aligned} \tag{3.1}$$

A nice introduction to Bézier curves can be found in [**311**, Chapter 2].

Here are some Bézier cubics that illustrate the *convex hull property*:



In general, $\sum_{i=0}^{n} \binom{n}{i}t^i(1-t)^{n-i} = 1$ implies that $\Phi(t)$ is a convex combination of $P_0, \dots, P_n$ and hence lies in the *convex hull* $\mathrm{Conv}(P_0, \dots, P_n) \subseteq \mathbb{R}^\ell$.

Tangent Lines and Curvature. A designer selects control points $P_0, \dots, P_n$ to create a Bézier curve. A nice feature of these curves is that it is easy to understand the tangent vectors and curvature of Bézier curve at the endpoints P_0 and P_n.

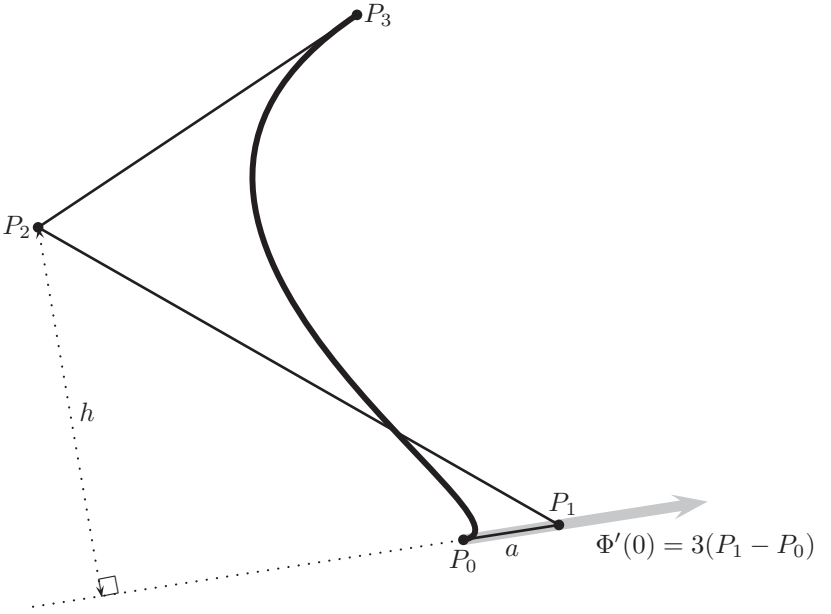
For simplicity, we focus on P_0, where $t = 0$. An easy calculation using (3.1) shows that the tangent vector at P_0 is

$$\Phi'(0) = n(P_1 - P_0).$$

[1] De Casteljau was first but not allowed to publish until six years after Bézier's 1968 paper.

When $n = 3$, we get a Bézier cubic. Here is a picture:

(3.2)



This curve has control points P_0, P_1, P_2, P_3 linked by line segments. The tangent vector $\Phi'(0) = 3(P_1 - P_0)$ is shown in gray.

The dotted lines in the picture relate to curvature. Let

$$a = \text{distance from } P_1 \text{ to } P_0 = \|P_1 - P_0\|$$
$$h = \text{distance from } P_2 \text{ to the line } \overleftrightarrow{P_0P_1}.$$

Then one can prove that the curvature at P_0 is given by the formula

$$\kappa(0) = \frac{n-1}{n}\frac{h}{a^2}. \tag{3.3}$$

The Bézier cubic (3.2) has a lot of curvature at P_0. This is obvious from the picture and also from the formula since a is much smaller than h.

One nice feature of these formulas is that when $t = 0$, we have:

- The position is determined by P_0.
- The tangent vector is determined by P_0, P_1.
- The curvature is determined by P_0, P_1, P_2.

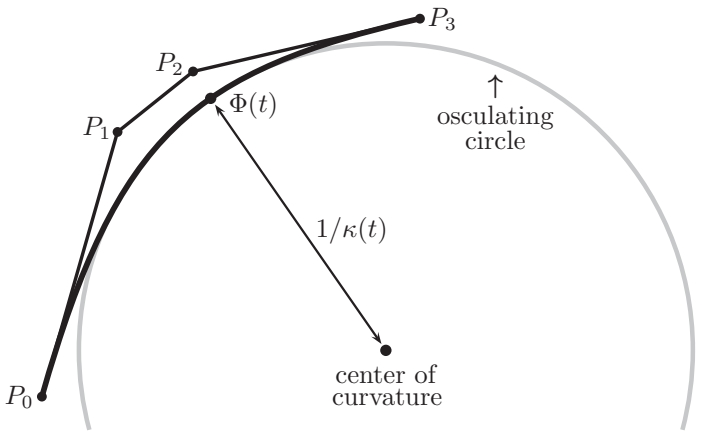
The position of the other control points has no effect on these quantitites.

There are analogous formulas for the tangent vector and curvature when $t = 1$ that depend only on P_{n-2}, P_{n-1}, P_n.

Visualizing Curvature. In differential geometry, curvature is often introduced via the *osculating circle*, which is the circle that best approximates the curve at a given point. Its radius is the reciprocal of the curvature and its center is the *center of curvature*. If $N(t)$ and $\kappa(t)$ are the unit normal and curvature at $\Phi(t)$, it follows that the center of curvature is

$$\Phi(t) + \frac{1}{\kappa(t)}N(t). \tag{3.4}$$

For a Bézier cubic, this is illustrated as follows:

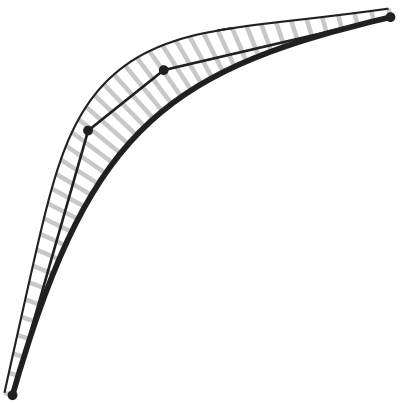


For a single point, this works nicely, but for multiple points, the osculating circles create a cluttered picture. It is simpler to focus on the centers of curvature, the *curve of centers of curvature*. However, its distance to original curve is the reciprocal of the curvature, making it harder to see how the curvature varies.

In geometric modeling, it is customary to replace the center of curvature (3.4) with the point

$$\Phi(t) - \kappa(t)N(t).$$

Thus we work on the other side of the curve and consider points whose distance to the curve is the curvature rather than its reciprocal. The resulting picture is the *curvature comb*:

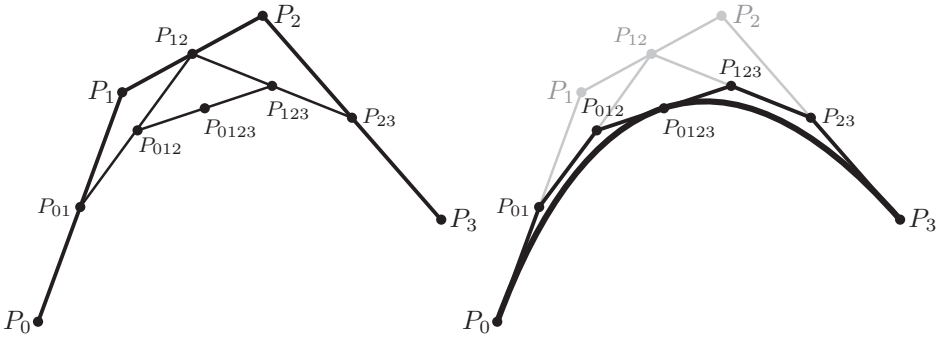


A web search of "AutoDesk Alias curvature comb" will lead to some nice images.

The de Casteljau Algorithm. The curvature comb is nice but requires knowing the curvature at many points. Right now, we only know the curvature at the endpoints. The de Casteljau algorithm gives us access to more points on a Bézier curve, as we illustrate for a Bézier cubic with control points P_0, P_1, P_2, P_3:

- Take successive midpoints to get P_{01}, P_{12}, P_{23}.
- Take successive midpoints of P_{01}, P_{12}, P_{23} to get P_{012}, P_{123}.
- Take the midpoint of P_{012}, P_{123} to get P_{0123}.

These successive midpoints are shown on the left of the following figure:



On the right, we have drawn the Bézier cubic for the control points P_0, P_1, P_2, P_3. The magic is that this curve breaks up into *two* Bézier cubics:

- The left half is the Bézier cubic with control points $P_0, P_{01}, P_{012}, P_{0123}$.
- The right half is the Bézier cubic with control points $P_{0123}, P_{123}, P_{23}, P_3$.

Since we know the curvature at the endpoints of a Bézier curve, we now know the curvature at P_{0123}. Iterating this process, we can find the curvature at many more points on the curve. Hence the curvature comb becomes easy to compute.

Two Variations on Bézier Curves. The Bézier curve (3.1) is defined on the interval $[0, 1]$ and features the binomial coefficients $\binom{n}{i}$. Both of these can be generalized:

- Change the interval. We can switch the interval from $[0, 1]$ to $[a, b]$ via

$$\Phi_{[a,b]}(t) = \frac{1}{(b-a)^n} \sum_{i=0}^{n} \binom{n}{i} (t-a)^i (b-t)^{n-i} P_i, \quad t \in [a, b].$$

  The convex hull property holds since $\sum_{i=0}^{n} \binom{n}{i} (t-a)^i (b-t)^{n-i} = (b-a)^n$.
- Use general weights. Instead of $\binom{n}{i}$, use positive weights w_i. This gives

$$\Phi_w(t) = \frac{1}{\sum_{i=0}^{n} w_i t^i (1-t)^{n-i}} \sum_{i=0}^{n} w_i t^i (1-t)^{n-i} P_i, \quad t \in [0, 1].$$

  The denominator guarantees that the convex hull property holds. For most weights, this is a rational parametrization. The denominator never vanishes since $t \in [0, 1]$ and the weights are positive.

It is also possible to combine these two variations. Here is an example that gives a preview of the connection to toric geometry.

EXAMPLE 3.1. Consider the interval $[0, n]$:

This is a one-dimensional lattice polytope with lattice points $0, \ldots, n$ and facets $0, n$. For positive weights $w_0, \ldots, w_n$ and control points $P_0, \ldots, P_n$, we get the weighted Bézier curve

$$\Phi(t) = \frac{1}{\sum_{i=0}^{n} w_i t^i (n-t)^{n-i}} \sum_{i=0}^{n} w_i t^i (n-t)^{n-i} P_i, \quad t \in [0, n].$$

The blending functions $t^i(n-t)^{n-i}$ are indexed by the lattice points $i = 0, \ldots, n$ and have two interesting properties:

- For $t \in [0, n]$, the functions t and $n - t$ give the distance from t to the facets $0, n$ of $[0, n]$,
- For each index i, the exponents i and $n - i$ give the distance from the lattice point $i \in [0, n]$ to the facets $0, n$.

Later in the section we will explain how this interpretation of $t^i(n-t)^{n-i}$ can be generalized to define blending functions indexed by the lattice points of an arbitrary lattice polytope. ◁▷

Font Design. There are many ways to design fonts such as Times Roman or Helvetica. In an *outline font*, one gives information that specifies the outline of each letter. Common examples include:

- TrueType fonts, which use quadratic Bézier curves.
- Type 1 PostScript fonts, which use cubic Bézier curves.

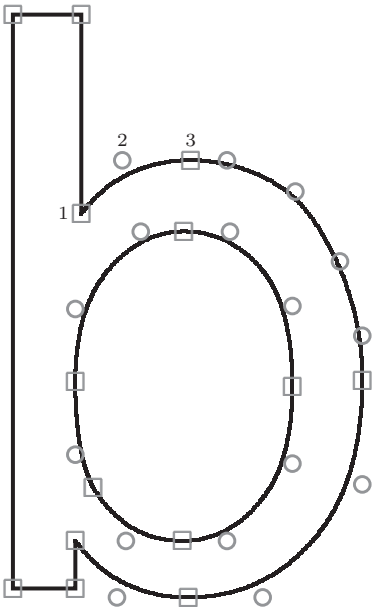
Example 3.2. In the TrueType font Monotype Arial, the letter "b" is built from line segments and quadratic Bézier curves of the form

$$\Phi(t) = (1-t)^2(a_0, b_0) + 2t(1-t)(a_1, b_1) + t^2(a_2, b_2), \quad t \in [0, 1].$$

This curve goes through $\square = (a_0, b_0), \square = (a_2, b_2)$, and $\bigcirc = (a_1, b_1)$ "pulls" the curve. In this notation, a Bézier quadratic has control points $\square, \bigcirc, \square$.

Now consider the whole letter:

(3.5)



The line segments are indicted by consecutive □'s, and the rest of the outline is composed of Bézier quadratics. For example, the points marked 1, 2, 3 in (3.5) are the control points $\square, \bigcirc, \square$ that define one of the Bézier quadratics.

The consecutive ○'s in the picture seem to indicate missing control points. When this happens, the control point is not shown because it is the *midpoint* of the the line segment connecting the consecutive ○'s. This midpoint convention allows the outline of the letter to be specified more efficiently. ◁▷

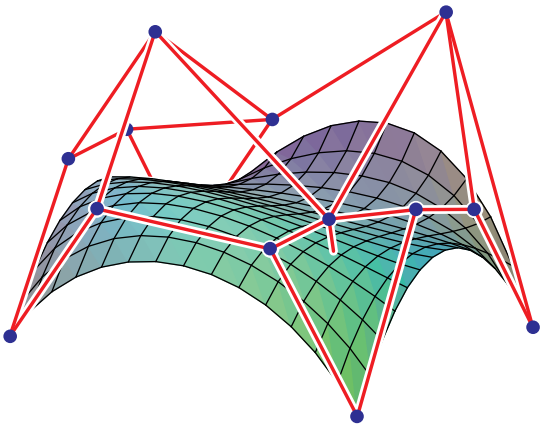
Bézier Surfaces. We begin with the most common Bézier surface, the *tensor product surface of degree* (m, n). Given control points $P_{ij} \in \mathbb{R}^\ell$, $0 \le i \le m$, $0 \le j \le n$, we get the parametrization

$$\Phi(s,t) = \sum_{i=0}^{m} \sum_{j=0}^{n} \binom{m}{i} \binom{n}{j} s^i (1-s)^{m-i} t^j (1-t)^{n-j} P_{ij}, \ (s,t) \in [0,1] \times [0,1]. \tag{3.6}$$

The image is called a *surface patch* since it is the portion of the surface parametrized by the unit square. Note that the convex hull property holds since the coefficients in (3.6) sum to 1.

Here is a splendid example of a bicubic Bézier patch (degree $(3, 3)$) taken from the paper [**160**]:

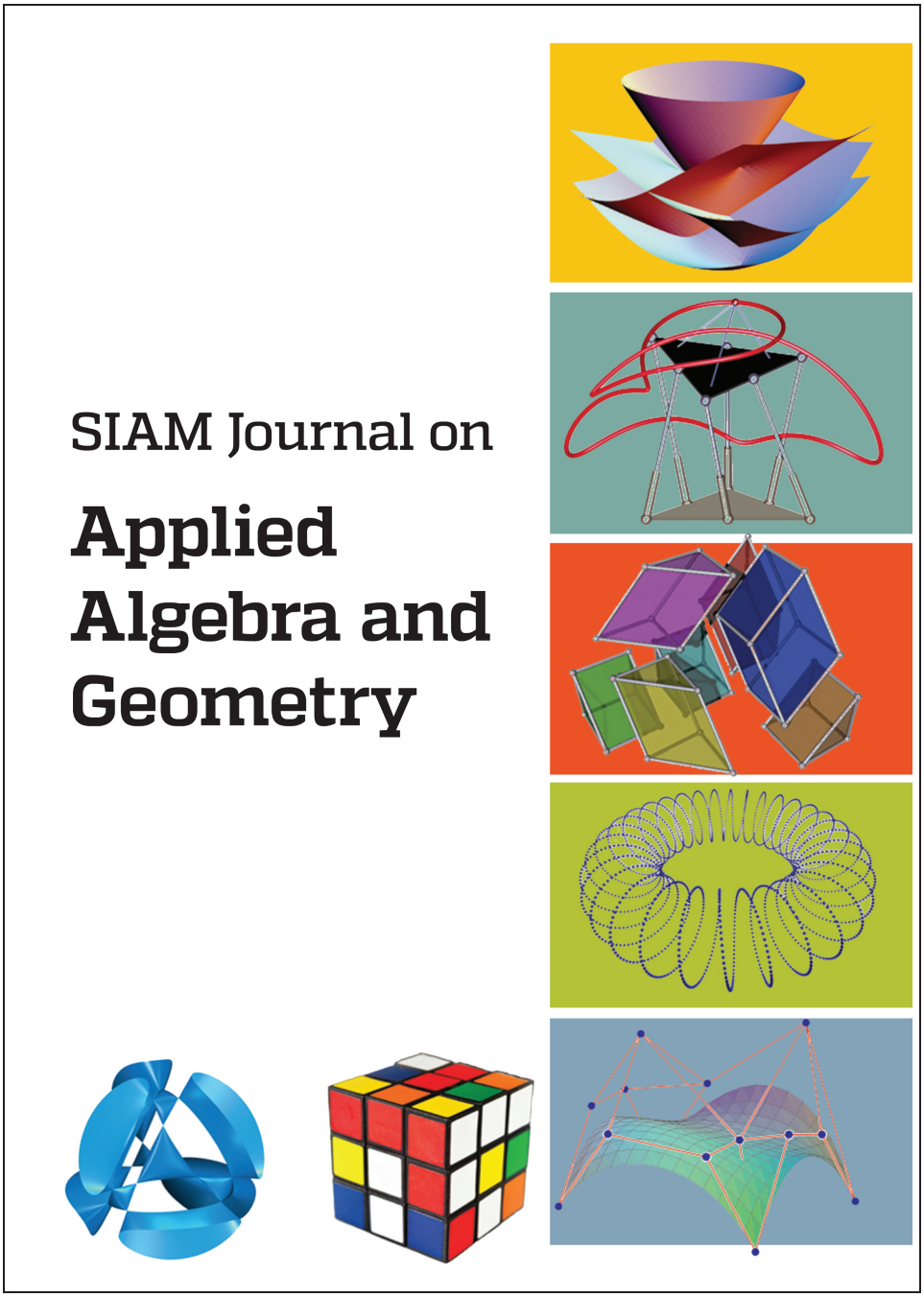
(3.7)



Note that the edges of this patch are Bézier cubics. Thirteen of the 16 control points P_{ij}, $0 \le i, j \le 3$, appear in (3.7). The red lines will be explained later in the section. This well-known image appears on the cover of the SIAM Journal on Applied Algebra and Geometry shown on page 95.

Bicubic patches are widely used to design automobiles, airplanes, and buildings. A nice example is the Guggenheim Bilbao:

This building was designed in CATIA, a program created by the French aerospace company Dassault Systems. Each titanium sheet is a Bézier patch. The book *Digital Gehry* [**261**] describes the process used to create this amazing structure.

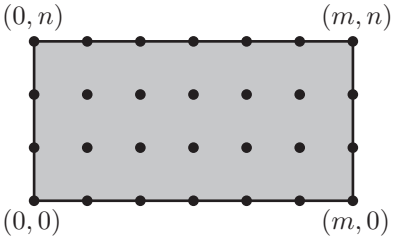

SIAM Journal on
Applied
Algebra and
Geometry

We can rescale the formula (3.6) to write a Bézier patch of degree (m, n) as

$$\Phi(s,t) = \frac{1}{m^m n^n} \sum_{i,j} \binom{m}{i}\binom{n}{j} s^i (m-s)^{m-i} t^j (n-t)^{n-j} P_{ij} \tag{3.8}$$

for (s, t) in the rectangle $P = [0, m] \times [0, n]$. The denominator $m^m n^n$ is needed for the convex hull property.

The sum in (3.8) is over the *lattice points* $P \cap \mathbb{Z}^2$ of P:

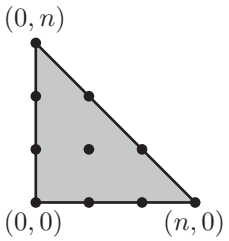


For each lattice point (i, j), the blending function $s^i(m-s)^{m-i}t^j(n-t)^{n-j}$ is built from the distance to the facets of P. We will soon generalize this to an arbitrary lattice polytope.

Triangular Bézier Patch of Degree n. Let $\Delta_2 = \mathrm{Conv}(0, e_1, e_2)$ be the standard simplex in $\mathbb{R}^2$. Given $n > 0$ and control points $P_{ij} \in \mathbb{R}^\ell$, $i, j \geq 0$, $i + j \leq n$, we have the triangular Bézier patch

$$\Phi(s,t) = \frac{1}{n^n} \sum_{i+j\leq n} \frac{n!}{i!j!(n-i-j)!} s^i t^j (n-s-t)^{n-i-j} P_{ij} \tag{3.9}$$

for (s, t) in the triangle $P = n\Delta_2$. As above, the n^n is needed for the convex hull property. The sum is again over the lattice points of P:



The blending function $s^i t^j (n-s-t)^{n-i-j}$ indexed by (i, j) is built out of distances from (s, t) and (i, j) to the facets, but for the facet opposite the origin, one needs to use the notion of *lattice distance*, to be defined in (3.10) below.

Weighted Surface Patches. The weighted versions of tensor product (3.6) and triangular (3.9) surface patches replace $\binom{m}{i}\binom{n}{j}$ and $\frac{n!}{i!j!(n-i-j)!}$ with positive weights w_{ij}. For example, the weighted patch defined on $[0, m] \times [0, n]$ is

$$\Phi(s,t) = \frac{1}{\sum_{i,j} w_{ij} s^i (m-s)^{m-i} t^j (n-t)^{n-j}} \sum_{i,j} w_{ij} s^i (m-s)^{m-i} t^j (n-t)^{n-j} P_{ij}.$$

The denominator ensures the convex hull property and doesn't vanish on P since the weights are positive. Similar to the curve case, weighted surface parametrizations typically involve rational functions rather than polynomials.

Lattice Polytopes. In 2002, Rimvydas Krasauskas [**238**] found a way to generalize the Bézier surface patches (3.6) and (3.9) to an arbitrary *lattice polytope* $P \subseteq \mathbb{R}^d$ of dimension d.

The polytope P is the convex hull of a finite set $\mathcal{A} \subseteq \mathbb{Z}^d$, though for us, the *facet presentation* of P is equally important. A *facet* $F \subseteq P$ is a codimension-one face whose supporting hyperplane is defined by

$$\langle m, n_F \rangle = -a_F, \quad -a_F \in \mathbb{Z},\ n_F \in \mathbb{Z}^n.$$

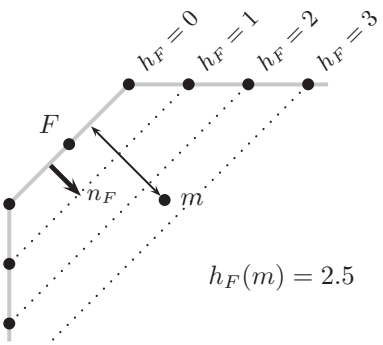
We specify this equation uniquely by requiring that $n_F \in \mathbb{Z}^n$ be primitive (meaning $n_F \notin m\mathbb{Z}^n$ for all integers $m \geq 2$) and inward-pointing. Then the desired facet presentation is

$$P = \{m \in \mathbb{R}^d \mid \langle m, n_F \rangle \geq -a_F \text{ for all facets } F\}.$$

Lattice Distance. For a point $m \in P$, the quantity

$$h_F(m) = \langle m, n_F \rangle + a_F \geq 0 \tag{3.10}$$

measures the number of parallel integral hyperplanes one must cross in order to get to the facet F. Here is an example of what this looks like in $\mathbb{R}^2$:



We call $h_F(m)$ the *lattice distance* from m to F. Note that $h_F(m)$ need not be the Euclidean distance. In this notation, the facet presentation becomes

$$P = \{m \in \mathbb{R}^n \mid h_F(m) \geq 0 \text{ for all facets } F\}.$$

Toric Bézier Functions. Fix a finite set $\mathcal{A} \subseteq \mathbb{Z}^d$ and let $P = \text{Conv}(\mathcal{A})$. For $\mathbf{a} \in \mathcal{A}$ and $m \in P$, define the *toric Bézier function*

$$\beta_{\mathbf{a}}^{\text{toric}}(m) = \prod_F h_F(m)^{h_F(\mathbf{a})}, \tag{3.11}$$

where the product is over all facets F of P. The exponents are nonnegative since $\mathbf{a} \in P$ and $\beta_{\mathbf{a}}^{\text{toric}}(m) \geq 0$ since $m \in P$. Here are some familiar examples:

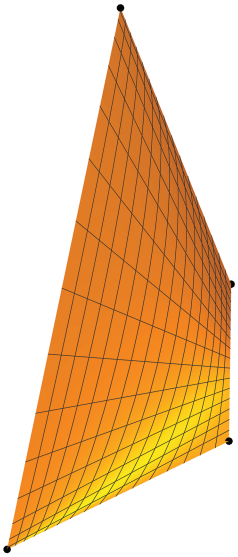
- $P = [0, n]$ gives the functions $t^i(n-t)^{n-i}$ used in the Bézier curve defined in Example 3.1.
- $P = [0, m] \times [0, n]$ gives the functions $s^i(m-s)^{m-i}t^j(n-t)^{n-j}$ used in the tensor product Bézier patch (3.8).
- $P = n\Delta_2$ gives the functions $s^i t^j (n-s-t)^{n-i-j}$ used in the triangular Bézier patch (3.9).

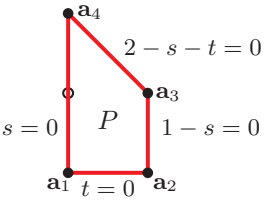
Toric Patches. The Bézier formulas (3.1) and (3.8) use binomial coefficients and (3.9) uses the trinomial coefficient $\frac{n!}{i!j!(n-i-j)!}$. Since these do not have an obvious toric generalization (more on this below), a *toric patch* uses positive weights $w = \{w_{\mathbf{a}}\}_{\mathbf{a}\in\mathcal{A}}$ and control points $\{P_{\mathbf{a}}\}_{\mathbf{a}\in\mathcal{A}} \subseteq \mathbb{R}^\ell$ to define $\Phi_{\mathcal{A},w} : P \to \mathbb{R}^\ell$ by

$$\Phi_{\mathcal{A},w}(m) = \frac{1}{\sum_{\mathbf{a}\in\mathcal{A}} w_{\mathbf{a}}\beta^{\text{toric}}_{\mathbf{a}}(m)} \sum_{\mathbf{a}\in\mathcal{A}} w_{\mathbf{a}}\beta^{\text{toric}}_{\mathbf{a}}(m)\, P_{\mathbf{a}}, \quad m \in P. \tag{3.12}$$

A common choice for $\mathcal{A}$ is the set of all lattice points in P. This is what happens for Bézier curves and surfaces. Here is an example where we use fewer.

EXAMPLE 3.3. Let $\mathcal{A} = \{\mathbf{a}_1 = (0,0), \mathbf{a}_2 = (1,0), \mathbf{a}_3 = (1,1), \mathbf{a}_4 = (0,2)\} \subseteq \mathbb{Z}^2$. Notice that $\mathcal{A}$ does not use all lattice points in $P = \text{Conv}(\mathcal{A})$.





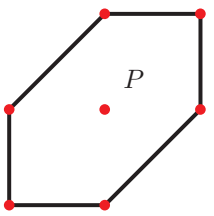
control points:
$P_1 = (0,0,0)$
$P_2 = (.5,1,0)$
$P_3 = (0,2,0)$
$P_4 = (0,1,2)$
weights: $1,1,1,1$

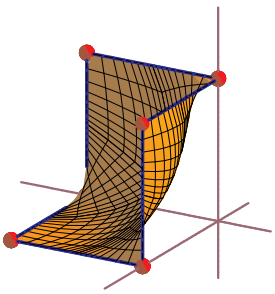
toric Bézier functions:

$$\beta^{\text{toric}}_1(s,t) = s^0t^0(1-s)^1(2-s-t)^2$$
$$\beta^{\text{toric}}_2(s,t) = s^1t^0(1-s)^0(2-s-t)^1$$
$$\beta^{\text{toric}}_3(s,t) = s^1t^1(1-s)^0(2-s-t)^0$$
$$\beta^{\text{toric}}_4(s,t) = s^0t^2(1-s)^1(2-s-t)^0$$

◁▷

EXAMPLE 3.4. Here is an example from [**327**] that uses all lattice points:



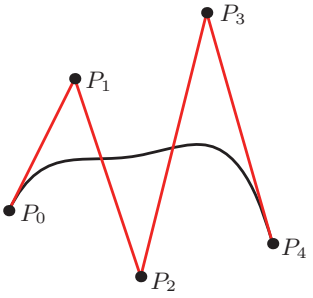


The "bulge" in the surface is governed by the control point indicated here ↑. ◁▷

Degenerations of Toric Patches. By letting the weights go to ∞, we can degenerate a toric patch into simpler patches. To illustrate how this works, we

begin with curves. A Bézier curve with standard weights $w_i = \binom{n}{i}$ and control points $P_0, \ldots, P_n$ goes through P_0 and P_n and is "pulled" toward $P_1, \ldots, P_{n-1}$.

EXAMPLE 3.5. Here is a Bézier quartic with standard weights $1, 4, 6, 4, 1$:



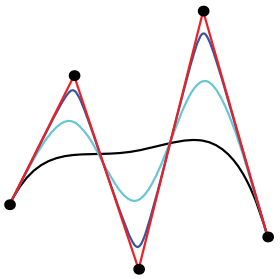
We can increase the "pull" of P_1, P_2, P_3 by varying the weights. But there are many ways to do this. If we use the weights

$$1, 4w^\alpha, 6w^\beta, 4w^\gamma, 1, \tag{3.13}$$

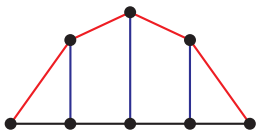
then we get the weighted Bézier quartic

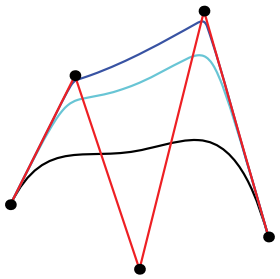
$$\Phi(t) = \frac{(1-t)^4 P_0 + 4w^\alpha t(1-t)^3 P_1 + 6w^\beta t^2(1-t)^2 P_2 + 4w^\gamma t^3(1-t) P_3 + t^4 P_4}{(1-t)^4 + 4w^\alpha t(1-t)^3 + 6w^\beta t^2(1-t)^2 + 4w^\gamma t^3(1-t) + t^4}$$

defined for $t \in [0, 1]$. When we let $w \to \infty$, the limit curve that depends on the exponents α, β, γ in (3.13). Each picture (a), (b), (c) below shows the original Bézier quartic in black and two weighted versions, one in light blue with bigger weights and one in dark blue with much bigger weights:

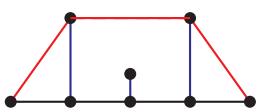


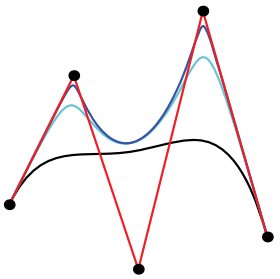
(a) weights: $1, 4w^3, 6w^4, 4w^3, 1$ exponent plot:



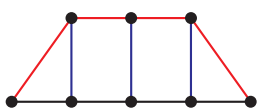


(b) weights: $1, 4w^3, 6w, 4w^3, 1$ exponent plot:





(c) weights: $1, 4w^3, 6w^3, 4w^3, 1$ exponent plot:



The *exponent plot* is explained as follows:

- The exponents of w in (3.13) are $0, \alpha, \beta, \gamma, 0$.

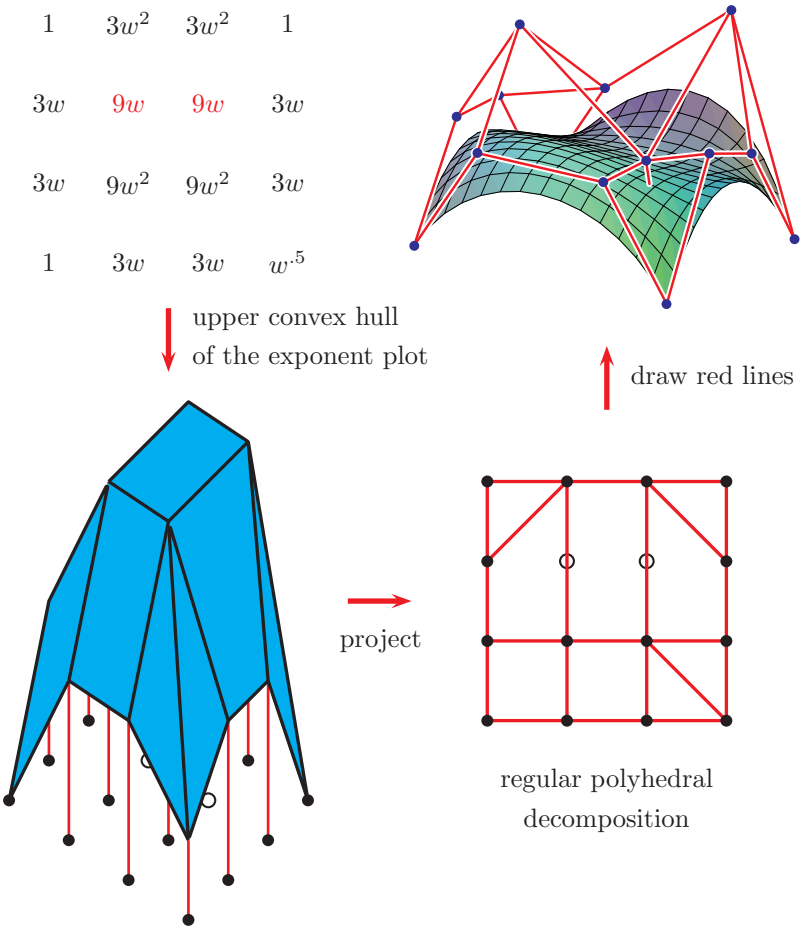
- The exponent plot consists of the points $(0,0),(1,\alpha),(2,\beta),(3,\gamma),(4,0)$.
- The *upper convex hull* of the exponent plot is outlined in red.

The limit curve as $w \to \infty$ is determined by the upper convex hull. The limit curve has four line segments in (a) and three in (b). In (c), the limit consists of two line segments and a weighted Bézier quadratic since the top of the upper convex hull is a copy of $[0,2]$. You can see the limiting behavior in the above pictures. ◁▷

EXAMPLE 3.6. We now explain the red lines in bicubic surface patch (3.7). For $(n,m) = (3,3)$, the products of binomial coefficients in (3.7) form a square:

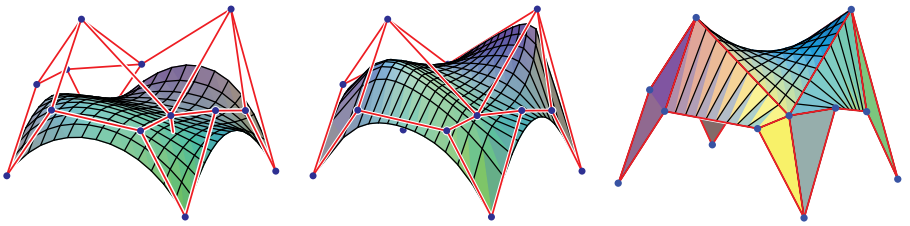
$$\begin{array}{cccc} 1 & 3 & 3 & 1 \\ 3 & 9 & 9 & 3 \\ 3 & 9 & 9 & 3 \\ 1 & 3 & 3 & 1 \end{array}$$

The idea is to introduce weights depending on a parameter w as in Example 3.5 and use the upper convex hull of the exponent plot to understand the limit as $w \to \infty$. We will consider the weights used in [**160**]:



The two weights shown in red do not contribute to the upper convex hull and, as indicated by the open circles, are not part of the regular polyhedral decomposition of $P = [0,3] \times [0,3]$ induced by the upper convex hull. The red lines in the decomposition of P become the red lines we originally saw in (3.7). This also explains why we see only 13 of the 16 control points: one missing point is hidden behind the surface, and the other two come from the open circles not used in the decomposition.

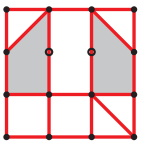
Furthermore, as shown in [**160**], letting $w \to \infty$ leads to a limiting configuration supported on the red lines:



$w = 1$ $w = 6$ $w = \infty$

The limit consists of 9 surface patches:

- 3 tensor-product Bézier patches.
- 4 triangular Bézier patches.
- 2 toric patches (shaded) that are *not* Bézier!



This example shows how toric patches can arise in a purely Bézier setting. ◁▷

The full story of Examples 3.5 and 3.6 is explained in Theorem 3 of [**160**].

Bernstein Polynomials. The polynomials $\binom{n}{i} t^i (1-t)^{n-i}$ appearing in the Bézier curve (3.1) where introduced in 1912 by Sergei Bernstein in his elegant proof of the Weierstrass Approximation Theorem [**26**]. To approximate a continuous function $F : [0,1] \to \mathbb{R}$ by a polynomial of degree n, Bernstein took a probabilistic approach. A player flips a biased coin where the probability of heads is t and tails is $1-t$, for some $t \in [0,1]$. Bernstein assumes the player flips the coin n times and is paid the sum $F(\frac{i}{n})$ if heads occurs exactly i times. Under these conditions, the expectation of the player is

$$E_n = \sum_{i=0}^{n} F\left(\tfrac{i}{n}\right) \binom{n}{i} t^i (1-t)^{n-i}. \tag{3.14}$$

Bernstein proves that $E_n \to F$ uniformly for $t \in [0,1]$.

The Bernstein polynomials $\binom{n}{i} t^i (1-t)^{n-i}$ are well known in numerical analysis and have many nice properties, as explained in [**124**, Theorem 7.3.1]. A modern version of Bernstein's result can be found in [**307**, Theorem 5.3.2].

Bernstein's proof of the approximation theorem extends easily to continuous functions $F : [0,1] \to \mathbb{R}^\ell$. Here, the graph of F is approximated by the Bézier curve $\Phi(t) = \sum_{i=0}^{n} \binom{n}{i} t^i (1-t)^{n-i} P_i$, where the i^{th} control point P_i is chosen to be the point $F(\frac{i}{n})$ on the curve. This explains why Bernstein polynomials appear naturally in Bézier curves. In approximation theory, n is usually large. The insight

of de Casteljau and Bézier is that in the geometric setting, small values of n lead to useful curves.

Linear Precision and Weights. One nice feature of Bernstein polynomials is that for an affine linear function $F : [0,1] \to \mathbb{R}^\ell$, the approximation is exact, i.e.,

$$F(x) = \sum_{i=0}^{n} \tbinom{n}{i} t^i (1-t)^{n-i} F\big(\tfrac{i}{n}\big).$$

This property is called *linear precision*.

For Bézier curves, linear precision helps explain the role of the weights $w_i > 0$ that appear in a weighted Bézier curve

$$\Phi(t) = \frac{1}{\sum_{i=0}^{n} w_i\, t^i (1-t)^{n-i}} \sum_{i=0}^{n} w_i\, t^i (1-t)^{n-i} P_i.$$

The convex hull property implies that if the P_i are collinear, then Φ parametrizes a line segment, though the parametrization usually has degree n. However, if $w_i = \binom{n}{i}$ and $P_0, \ldots, P_n$ are collinear and equally spaced, then linear precision implies that Φ is affine linear. A lot of cancellation is needed for the degree to drop from n to 1. Binomial coefficients are indeed special.

Linear Precision for General Patches. Fix a finite set $\mathcal{A} \subseteq \mathbb{Z}^d$ and assume that $P = \mathrm{Conv}(\mathcal{A}) \subseteq \mathbb{R}^d$ has dimension d. A *general patch* has two ingredients:

- Continuous blending functions $\hat\beta_{\mathbf{a}} : P \to \mathbb{R}_{\geq 0}$, $\mathbf{a} \in \mathcal{A}$, with $\sum_{\mathbf{a}\in\mathcal{A}} \hat\beta_{\mathbf{a}} = 1$.
- Control points $P_{\mathbf{a}} \in \mathbb{R}^\ell$ for $\mathbf{a} \in \mathcal{A}$.

Then the patch is the function $\Phi : P \to \mathbb{R}^\ell$ defined by

$$\Phi(m) = \sum_{\mathbf{a}\in\mathcal{A}} \hat\beta_{\mathbf{a}}(m)\, P_{\mathbf{a}}, \quad m \in P. \tag{3.15}$$

Note that general patches satisfy the convex hull property.

Example 3.7. A toric patch (3.12) can be written

$$\Phi_{\mathcal{A},w}(m) = \sum_{\mathbf{a}\in\mathcal{A}} \hat\beta^{\mathrm{toric}}_{\mathbf{a},w}(m)\, P_{\mathbf{a}},$$

where the *toric blending function* $\hat\beta^{\mathrm{toric}}_{\mathbf{a},w}$ is the rational function

$$\hat\beta^{\mathrm{toric}}_{\mathbf{a},w}(m) = \frac{w_{\mathbf{a}}\beta^{\mathrm{toric}}_{\mathbf{a}}(m)}{\sum_{\mathbf{a}\in\mathcal{A}} w_{\mathbf{a}}\beta^{\mathrm{toric}}_{\mathbf{a}}(m)}$$

and $\beta^{\mathrm{toric}}_{\mathbf{a}}$ is the toric Bézier function defined in (3.11). ◁▷

Example 3.8. For blending functions $\{\hat\beta_{\mathbf{a}}\}_{\mathbf{a}\in\mathcal{A}}$ as above, the canonical choice $P_{\mathbf{a}} = \mathbf{a}$ of control points gives the *tautological patch* defined by

$$\Phi^{\mathrm{taut}}(m) = \sum_{\mathbf{a}\in\mathcal{A}} \hat\beta_{\mathbf{a}}(m)\, \mathbf{a}.$$

Note that $\Phi^{\mathrm{taut}}(m) \in P$ by the convex hull property. Thus $\Phi^{\mathrm{taut}} : P \to P$. ◁▷

For a general patch (3.15), one can prove that the following are equivalent:

- For any affine linear function $\Lambda : \mathbb{R}^d \to \mathbb{R}^\ell$, we have

$$\Lambda(m) = \sum_{\mathbf{a}\in\mathcal{A}} \hat\beta_{\mathbf{a}}(m)\, \Lambda(\mathbf{a}), \quad m \in P.$$

- $m = \sum_{\mathbf{a}\in\mathcal{A}} \hat{\beta}_{\mathbf{a}}(m)\,\mathbf{a}$ for all $m \in P$. In other words, the tautological map $\Phi^{\text{taut}} : P \to P$ is the identity.

Patches satisfying these conditions have *linear precision*. The paper [**159**] studies linear precision in geometric modeling.

Toric Patches and Strict Linear Precision. Bézier curves (3.1), tensor-product patches (3.6), and triangular patches (3.9) have linear precision. These generalize to *Bézier simploids*, which are toric patches built from products of simplices with coefficients given by products of multinomial coefficients. See [**159**, Example 3.8] for a precise definition. It is well known that Bézier simploids have linear precision.

DEFINITION 3.9. A finite set $\mathcal{A} \subseteq \mathbb{Z}^n$ has *strict linear precision* for positive weights $w = \{w_{\mathbf{a}}\}_{\mathbf{a}\in\mathcal{A}}$ if the toric patch $\Phi_{\mathcal{A},w}$ (3.12) has linear precision, i.e., if

$$m = \frac{1}{\sum_{\mathbf{a}\in\mathcal{A}} w_{\mathbf{a}}\beta^{\text{toric}}_{\mathbf{a}}(m)} \sum_{\mathbf{a}\in\mathcal{A}} w_{\mathbf{a}}\beta^{\text{toric}}_{\mathbf{a}}(m)\,\mathbf{a} = \sum_{\mathbf{a}\in\mathcal{A}} \hat{\beta}^{\text{toric}}_{\mathbf{a},w}(m)\,\mathbf{a}$$

for all $m \in P$.

It is likely that Bézier simploids are the only toric patches with strict linear precision. Here is recent result from [**99**].

THEOREM 3.10. *Let $\mathcal{A}$ be the set of lattice points in a d-dimensional lattice polytope $P \subseteq \mathbb{R}^d$. Then $\mathcal{A}$ has strict linear precision for positive weights $w = \{w_{\mathbf{a}}\}_{\mathbf{a}\in\mathcal{A}}$ if and only if*

(1) *The sum of the facet normals n_F of P is equal to zero.*
(2) *The polynomial $\sum_{\mathbf{a}\in\mathcal{A}} w_{\mathbf{a}}\beta^{\text{toric}}_{\mathbf{a}}$ is constant.*

The polynomial in (2) is the denominator of the toric patch $\Phi_{\mathcal{A},w}$. It follows that toric patches with strict linear precision are polynomial parametrizations.

The proof of Theorem 3.10 has a surprise—it uses ideas from algebraic statistics, including the concept of *maximum likelihood degree* from Section 2.2.

Algebraic Statistics. The toric blending functions $\hat{\beta}^{\text{toric}}_{\mathbf{a},w}$ defined in Example 3.7 give a map

$$\begin{aligned} P &\longrightarrow \mathbb{R}^{\mathcal{A}}_{\geq 0} \\ m &\longmapsto \big(\hat{\beta}^{\text{toric}}_{\mathbf{a},w}(m)\big)_{\mathbf{a}\in\mathcal{A}} \end{aligned} \tag{3.16}$$

whose image is denoted $Y_{\mathcal{A},w}$. Points of $Y_{\mathcal{A},w}$ have non-negative coordinates that sum to 1, so that $Y_{\mathcal{A},w} \subseteq \mathbb{R}^{\mathcal{A}}_{\geq 0}$ is a statistical model. Example 3.5 of [**159**] implies that $Y_{\mathcal{A},w}$ is a what statisticians call a *log-linear model*, though these days the term *toric model* is also used.

In Chapter 2 we encountered the notions of *maximum likelihood estimate* and *maximum likelihood degree* from algebraic statistics. The phylogenetic model (2.25) discussed in Example 2.17 has ML degree equal to 92. But in some special cases, a statistical model has ML degree one, which is nice because this implies that the ML estimate is a rational function of the data.

In 2014 June Huh [**215**] proved a powerful result about varieties in $(\mathbb{C}^*)^n$ of ML degree one. This is relevant because, as we explain below, strict linear precision implies that $Y_{\mathcal{A},w}$ has ML degree one. Huh's result plays a central role in the proof of Theorem 3.10 given in [**99**, Theorem 7.2].

At first glance, the relation between algebraic statistics and geometric modeling seems unexpected. But keep in mind that statistical thinking inspired Bernstein in 1912 to define the polynomials that appear so naturally in geometric modeling.

We now explore one further aspect of this connection.

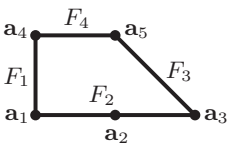
Toric Patches and Rational Linear Precision. The definition of strict linear precision is very restrictive since it uses the toric blending functions $\hat{\beta}^{\mathsf{toric}}_{\mathbf{a},w}$ from Example 3.7. Following [**159**], we say $\mathcal{A}$ has *rational linear precision* for positive weights $w = \{w_{\mathbf{a}}\}_{\mathbf{a}\in\mathcal{A}}$ if there are rational blending functions $\hat{\beta}_{\mathbf{a}} : P \to \mathbb{R}_{\geq 0}$ for $\mathbf{a} \in \mathcal{A}$ satsifying $\sum_{\mathbf{a}\in\mathcal{A}} \hat{\beta}_{\mathbf{a}} = 1$ with the following two properties:

- The $\hat{\beta}_{\mathbf{a}}$ have linear precision, i.e., for all $m \in P$, we have
$$m = \sum_{\mathbf{a}\in\mathcal{A}} \hat{\beta}_{\mathbf{a}}(m)\,\mathbf{a}.$$
- The image of the map
$$m \longmapsto \big(\hat{\beta}_{\mathbf{a}}(m)\big)_{\mathbf{a}\in\mathcal{A}}$$
is the statistical model $Y_{\mathcal{A},w}$ defined in (3.16).

In other words, rational linear precision for $\mathcal{A}, w$ means that we can parametrize $Y_{\mathcal{A},w}$ by rational blending functions that have linear precision.

Note that strict linear precision implies rational linear precision since the toric blending functions satisfy the above condition in the strict case. The converse fails, as shown by the following example.

EXAMPLE 3.11. Consider the trapezoid



with $\mathcal{A} = \{\mathbf{a}_1 = (0,0), \mathbf{a}_2 = (1,0), \mathbf{a}_3 = (2,0), \mathbf{a}_4 = (0,1), \mathbf{a}_5 = (1,1)\} \subseteq \mathbb{Z}^2$.

Let h_i denote the lattice distance to the facet F_i. For weights $w = (1,2,1,1,1)$ and control points $P_1, \ldots, P_5$, we get the toric patch

$$m \longmapsto \frac{1}{(2-t)(2-2t-t^2)}\big(h_3^2h_4P_1+2h_1h_3h_4P_2+h_1^2h_4P_3+h_2h_3P_4+h_1h_2P_5\big). \tag{3.17}$$

Using Theorem 3.10, we see that strict linear precision fails since the facet normals do not sum to zero. Also, the denominator in the above formula is nonconstant.

However, one can show that for the same weights $w = (1,2,1,1,1)$, $\mathcal{A}$ has rational linear precision with the blending functions

$$m \longmapsto \frac{h_3^2h_4}{(2-t)^2}P_1 + \frac{2h_1h_3h_4}{(2-t)^2}P_2 + \frac{h_1^2h_4}{(2-t)^2}P_3 + \frac{h_2h_3}{2-t}P_4 + \frac{h_1h_2}{2-t}P_5. \tag{3.18}$$

Formulas (3.17) and (3.18) are similar, but there is no theory that explains their relation. It is also unknown which polytopes have rational linear precision. ◁▷

We close with the lovely connection between rational linear precision and ML degree one proved in [**159**, Proposition 4.1].

THEOREM 3.12. *$\mathcal{A}$ has rational linear precision for w if and only if $Y_{\mathcal{A},w}$ has ML degree one, and when this happens, the ML estimate gives a parametrization with linear precision.*

There is a *lot* more to say about this subject, including many open questions. But for lack of space, we must stop here. You can consult the papers [**99**] and [**159**] and the references therein for further reading.

Final Comments. In spite of many missing details, we hope to have convinced you that geometric modeling has compelling pictures, wonderful applications, and an unexpected connection to algebraic statistics.

It is now time for the algebra of geometric modeling to take center stage.

3.2. Algebra of Modeling

We begin with curves. In the previous section, a weighted Bézier curve in the plane was written

$$t \in [0,1] \longmapsto \frac{1}{\sum_{i=0}^n w_i t^i (1-t)^{n-i}} \sum_{i=0}^n w_i t^i (1-t)^{n-i} P_i \in \mathbb{R}^2$$

for control points $P_0, \dots, P_n \in \mathbb{R}^2$. The algebraic structure of this situation is best revealed by working projectively over the algebraically closed field $\mathbb{C}$. Since $P_i = (\alpha_i, \beta_i) \in \mathbb{R}^2$ gives $(\alpha_i, \beta_i, 1) \in \mathbb{P}^2_{\mathbb{R}} \subseteq \mathbb{P}^2_{\mathbb{C}} = \mathbb{P}^2$, we can homogenize to obtain

$$(s,t) \in \mathbb{P}^1 \longmapsto \Big(\sum_{i=0}^n w_i \alpha_i t^i (s-t)^{n-i}, \sum_{i=0}^n w_i \beta_i t^i (s-t)^{n-i}, \sum_{i=0}^n w_i t^i (s-t)^{n-i} \Big) \in \mathbb{P}^2.$$

Each entry is a homogeneous polynomial of degree n in s, t.

More generally, a triple of homogeneous polynomials $a, b, c \in k[s,t]$ of degree n with $\gcd(a,b,c) = 1$ gives a map

$$\Phi : \mathbb{P}^1 \longrightarrow \mathbb{P}^2, \quad \Phi(s,t) = (a(s,t), b(s,t), c(s,t)), \tag{3.19}$$

whose image is a curve $C \subseteq \mathbb{P}^2$. Here, $k = \mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$. The gcd condition guarantees that a, b, c never vanish simultaneously on $\mathbb{P}^1$, so that Φ is defined on all of $\mathbb{P}^1$.

Algebraically, (3.19) gives the homogenous ideal

$$I = \langle a, b, c \rangle \subseteq k[s,t]. \tag{3.20}$$

Commutative algebra has developed a rich variety of tools to study homogeneous ideals in graded rings such as $k[s,t]$. For the ideal (3.20), we will see that *syzygies* and the *Rees algebra* relate nicely to concepts discovered by the geometric modeling community.

The Degree Formula. The parametrization (3.19) involves three types of degree:

- $n = \deg(a) = \deg(b) = \deg(c)$ is the degree of the polynomials $a, b, c \in k[s,t]$.
- $\deg(C)$ is the degree of the irreducible homogeneous polynomial $F \in k[x,y,z]$ with $C = \mathbf{V}(F)$.
- $\deg(\Phi)$ is the number of points of $\mathbb{P}^1$ that map via Φ to a generic point of C when working over $\mathbb{C}$.

These numbers are related by the following well-known formula.

THEOREM 3.13. *For $\Phi : \mathbb{P}^1 \to \mathbb{P}^2$ as in (3.19), we have*

$$n = \deg(\Phi) \cdot \deg(C).$$

EXAMPLE 3.14. $\Phi(s,t) = (s^2+t^2, 2s^2, s^2-t^2)$ is generically 2:1 and parametrizes the line $y = x + z$. Thus $n = 2$, $\deg(\Phi) = 2$ and $\deg(C) = 1$, and Theorem 3.13 becomes $2 = 2 \cdot 1$. ◁▷

Implicitization. We call $F = 0$ the *implicit equation* of C, and *implicitization* is the process of going from the parametrization $\Phi = (a, b, c)$ to the implicit equation $F = 0$. This is a problem in elimination theory since eliminating s, t from the equations

$$x - a(s,t) = y - b(s,t) = z - c(s,t) = 0$$

gives the implicit equation $F(x, y, z) = 0$.

In Section 1.2, we learned that elimination can be done using resultants or Gröbner bases. In particular, implicitization via Gröbner bases is a wonderful topic. Interested readers should consult [**104**, Chapter 3] for the basic theory and [**2**] for some recent developments. But for curves, the *syzygies* of the ideal (3.20) give a lovely way to think about implicitization that combines algebra and geometry.

Syzygies. By the Hilbert Syzygy Theorem, the ideal $I = \langle a, b, c\rangle \subseteq R = k[s, t]$ in (3.20) has a free resolution

$$0 \longrightarrow R(-n-\mu_1) \oplus R(-n-\mu_2) \xrightarrow{A} R(-n)^3 \xrightarrow{[a,b,c]} I \longrightarrow 0 \tag{3.21}$$

with $\mu_1 + \mu_2 = n$. This follows since a, b, c are relatively prime of degree n. A proof can be found in [**108**].

Elements $f_1, \ldots, f_d$ in a commutative ring R have a *syzygy module*

$$\operatorname{Syz}(f_1, \ldots, f_d) = \Big\{(h_1, \ldots, h_d) \in R^d \;\Big|\; \sum_{i=1}^{d} h_i f_i = 0\Big\}.$$

In general, syzygy modules can have a very complicated structure. The exact sequence (3.21) tells us that the syzygy module $\operatorname{Syz}(a, b, c)$ is free, with generators p, q of degrees

$$\deg(p) = \mu = \min(\mu_1, \mu_2) \le \max(\mu_1, \mu_2) = n - \mu = \deg(q).$$

It is customary to call $\{p, q\}$ a *μ-basis* of the syzygy module.

Syzygy is an astronomical term that refers to the alignment of three celestial bodies in a straight line. In 1853, Sylvester published *On a Theory of the Syzygetic Relations* [**341**]. This long memoir defines many colorful terms, including the Bezoutians and Bezoutics mentioned in Section 1.1. To help the reader, Sylvester added a *Glossary of New or Unusual Terms*, where he defines a *syzygetic function* to be a sum of given polynomials "with arbitrary functional multipliers, which are termed the syzygetic multipliers". He goes on to say

> When a syzygetic function of a given set of functions can be made to vanish, they are said to be syzygetically related.

Thus, what we now call a *syzygy on* $f_1, \ldots, f_d$ consists of the syzygetic multipliers $h_1, \ldots, h_d$ that give a syzygetic relation $\sum_{i=1}^{d} h_i f_i = 0$.

The exact sequence (3.21) was proved by Meyer in 1887 [**273**], who conjectured that a similar result should hold for any number of homogeneous generators. Hilbert proved the Syzygy Theorem in 1890 [**205**], though he never used the term "syzygy". Hilbert's first application of his theorem was a proof of Meyer's conjecture.

Moving Lines. A geometric approach to syzygyies was introduced in the 1995 paper *Implicitization of using moving curves and surfaces* by Sederberg and

Chen [**314**]. In the affine setting, the parametrization is written

$$x = \frac{a(t)}{c(t)}, \quad y = \frac{b(t)}{c(t)},$$

which we regard as a moving point in the plane. Then a *moving line* is an equation $A(t)x + B(t)y + C(t) = 0$ that depends polynomially on the parameter t, and the moving line *follows* the parametrization if the moving point lies on the moving line for all values of t. After clearing denominators, this gives the equation

$$A(t)a(t) + B(t)b(t) + C(t)c(t) = 0,$$

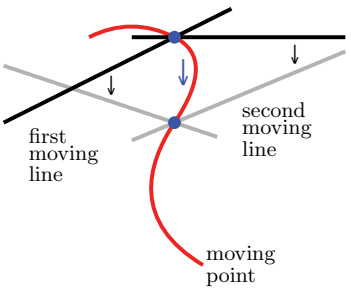
which tells us that we are in a syzygy module. This picture translates easily to the projective setting, with the result that a moving line $Ax + By + Cz = 0$ in $\mathbb{P}^2$ follows the parametrization $\Phi = (a, b, c) : \mathbb{P}^1 \to \mathbb{P}^2$ if and only if

$$(A, B, C) \in \operatorname{Syz}(a, b, c).$$

Here, $a, b, c, A, B, C \in k[s, t]$ are homogenous polynomials, where a, b, c have degree n and A, B, C have degree m for some m, reflecting that fact that $\operatorname{Syz}(a, b, c)$ is a graded module.

Intersection of Two Moving Lines. The point of intersection of two moving lines is a moving point:

(3.22)



This picture suggests that the parametrization $\Phi = (a, b, c)$ should come from the intersection of two moving lines in $\mathbb{P}^2$. Let's apply this idea to a μ-basis $\{p, q\}$, which we write as

$$p_1x + p_2y + p_3z = 0 \text{ and } q_1x + q_2y + q_3z = 0.$$

Since both go through the point (a, b, c), we get the syzygies

$$p_1a + p_2b + p_3c = 0 \text{ and } q_1a + q_2b + q_3c = 0. \tag{3.23}$$

By properties of determinants, we also have the equations

$$p_1 \det\begin{bmatrix} p_2 & q_2 \\ p_3 & q_3 \end{bmatrix} - p_2 \det\begin{bmatrix} p_1 & q_1 \\ p_3 & q_3 \end{bmatrix} + p_3 \det\begin{bmatrix} p_1 & q_1 \\ p_2 & q_2 \end{bmatrix} = 0$$

$$q_1 \det\begin{bmatrix} p_2 & q_2 \\ p_3 & q_3 \end{bmatrix} - q_2 \det\begin{bmatrix} p_1 & q_1 \\ p_3 & q_3 \end{bmatrix} + q_3 \det\begin{bmatrix} p_1 & q_1 \\ p_2 & q_2 \end{bmatrix} = 0.$$

The determinants have degree $\mu + (n - \mu) = n$, same as a, b, c, so (3.23) implies that up to a scalar multiple, the parametrization is given by

$$a = \det\begin{bmatrix} p_2 & q_2 \\ p_3 & q_3 \end{bmatrix}, \; b = -\det\begin{bmatrix} p_1 & q_1 \\ p_3 & q_3 \end{bmatrix}, \; c = \det\begin{bmatrix} p_1 & q_1 \\ p_2 & q_2 \end{bmatrix}. \tag{3.24}$$

Thus a, b, c are the 2×2 minors (with appropriate signs) of the matrix

$$A = \begin{bmatrix} p_1 & q_1 \\ p_2 & q_2 \\ p_3 & q_3 \end{bmatrix}$$

that appears in the exact sequence (3.21). This is the Hilbert-Burch Theorem!

It is fun to see how the picture (3.22) reveals the internal structure of the exact sequence (3.21) described by Hilbert-Burch. More interesting is that (3.22) was discovered by the geometric modeling community, which in the early 1990s was unaware of free resolutions and syzygy modules.

Implicitization via Moving Lines. We next present two related implicitization methods that use syzygies.

METHOD 1: MOVING LINES OF DEGREE $n-1$. A key observation is that there are n linearly independent moving lines of degree $n-1$. In terms of the syzygy module, this translates to the equality

$$\dim \operatorname{Syz}(a, b, c)_{n-1} = n. \tag{3.25}$$

We will prove this below using a μ-basis. Then let

$$(A_i, B_i, C_i) \in \operatorname{Syz}(a, b, c)_{n-1} \subseteq R^3_{n-1}, \ 1 \le i \le n,$$

be a basis, where as usual $R = k[s, t]$. This gives moving lines $A_i x + B_i y + C_i z$, which we rewrite using the monomials $s^{n-1}, s^{n-2}t, \ldots, st^{n-2}, t^{n-1}$ of degree $n-1$:

$$A_i x + B_i y + C_i z = \sum_{j=0}^{n-1} \ell_{ij}(x, y, z)\, s^{n-1-j} t^j.$$

We assemble the $\ell_{ij}(x, y, z)$ into an $n \times n$ matrix

$$M_{n-1} = \big(\ell_{ij}(x, y, z)\big)_{1 \le i \le n, 0 \le j \le n-1}$$

of linear forms. In 1995, Sederberg and Chen [**314**] proved the following result.

THEOREM 3.15. *Let $F = 0$ be the implicit equation of $C = \Phi(\mathbb{P}^1) \subseteq \mathbb{P}^2$. Then the $n \times n$ matrix of linear forms M_{n-1} defined above satisfies*

$$F^{\deg(\Phi)} = \det M_{n-1}.$$

We defer the proof until later. Note that the theorem is consistent with the degree formula in Theorem 3.13 since $\det M_{n-1}$ clearly has degree n.

An interesting feature of Theorem 3.15 is that it uses a matrix built from syzygies, just as Theorem 1.18 in Section 1.2 did surface implicitization using a matrix also built from syzygies.

METHOD 2: μ-BASES. The matrix M_{n-1} in Theorem 3.15 uses n linearly independent moving lines, yet the picture (3.22) suggests that we should need only two, namely those coming from a μ-basis of $\operatorname{Syz}(a, b, c)$. This intuition is correct, as we now explain.

Let the μ-basis be $p = (p_1, p_2, p_3)$ and $q = (q_1, q_2, q_3)$. By abuse of notation, we define the polynomials

$$p = p_1 x + p_2 y + p_3 z \text{ and } q = q_1 x + q_2 y + q_3 z.$$

Hence the moving lines are $p = 0$ and $q = 0$.

As polynomials, p and q are bihomogeneous of degrees μ and $n-\mu$ in s,t and degree 1 in x,y,z. To eliminate s,t, we use the Sylvester resultant from Section 1.2. Regarding p,q as homogeneous polynomials in s,t, the resultant $\operatorname{Res}(p,q)$ has degree

$$\deg_{s,t}(p)+\deg_{s,t}(q)=\mu+(n-\mu)=n$$

in x,y,z. In 1998, Cox, Sederberg and Chen [**108**] proved the following.

THEOREM 3.16. *Let $F=0$ be the implicit equation of $C=\Phi(\mathbb{P}^1)\subseteq\mathbb{P}^2$. Then a μ-basis $\{p,q\}$ of $\operatorname{Syz}(a,b,c)$ satisfies*

$$F^{\deg(\Phi)}=\operatorname{Res}(p,q).$$

PROOF. As we learned in Chapter 1, $\operatorname{Res}(p,q)$ vanishes at $(x_0,y_0,z_0)\in\mathbb{P}^2$ if and only if there exists $(s_0,t_0)\in\mathbb{P}^1$ such that

$$\begin{aligned}p_1(s_0,t_0)x_0+p_2(s_0,t_0)y_0+p_3(s_0,t_0)z_0&=0\\ q_1(s_0,t_0)x_0+q_2(s_0,t_0)y_0+q_3(s_0,t_0)z_0&=0.\end{aligned}$$

Arguing as in (3.23) and (3.24), these equations tell us that up to a scalar multiple,

$$\begin{aligned}x_0&=\det\begin{bmatrix}p_2(s_0,t_0) & q_2(s_0,t_0)\\ p_3(s_0,t_0) & q_3(s_0,t_0)\end{bmatrix}=a(s_0,t_0)\\ y_0&=-\det\begin{bmatrix}p_1(s_0,t_0) & q_1(s_0,t_0)\\ p_3(s_0,t_0) & q_3(s_0,t_0)\end{bmatrix}=b(s_0,t_0)\\ z_0&=\det\begin{bmatrix}p_1(s_0,t_0) & q_1(s_0,t_0)\\ p_2(s_0,t_0) & q_2(s_0,t_0)\end{bmatrix}=c(s_0,t_0).\end{aligned}$$

Thus $\operatorname{Res}(p,q)$ vanishes precisely on C and hence is a power of F, where $F=0$ is the implicit equation of C. We saw above that $\operatorname{Res}(p,q)$ has degree n, and then we are done since $n=\deg(\Phi)\cdot\deg(C)$ by Theorem 3.13. □

The μ-basis $\{p,q\}$ also relates nicely to METHOD 1. We assume $\mu\geq 1$ for simplicity. Since $\{p,q\}$ is a basis of the graded R-module $\operatorname{Syz}(a,b,c)$, $R=k[s,t]$, we have

$$\operatorname{Syz}(a,b,c)_{n-1}=R_{n-1-\mu}p\oplus R_{\mu-1}q.$$

It follows that the n syzygies

$$s^{n-1-\mu}p,\ s^{n-2-\mu}tp,\ \dots,\ t^{n-1-\mu}p,\ s^{\mu-1}q,\ s^{\mu-2}tq,\ \dots,\ t^{\mu-1}q \tag{3.26}$$

form a basis of $\operatorname{Syz}(a,b,c)_{n-1}$ as a vector space over k, proving (3.25). Furthermore, if we write

$$\begin{aligned}p&=a_0s^{\mu}+a_1s^{\mu-1}t+\cdots+a_\mu t^{\mu}\\ q&=b_0s^{n-\mu}+b_1s^{n-\mu-1}t+\cdots+b_{n-\mu}t^{n-\mu},\end{aligned}$$

where $a_0, \dots, a_\mu, b_0, \dots, b_{n-\mu}$ are linear in x, y, z, then expressing (3.26) in terms of the monomials of degree $n-1$ in s, t gives

$$\begin{bmatrix} s^{n-1-\mu}p \\ s^{n-2-\mu}tp \\ \vdots \\ t^{n-1-\mu}p \\ s^{\mu-1}q \\ s^{\mu-2}tq \\ \vdots \\ t^{\mu-1}q \end{bmatrix} = \begin{bmatrix} a_0 & a_1 & & \cdots & & a_\mu & & & \\ & a_0 & a_1 & & \cdots & & a_\mu & & \\ & & \ddots & \ddots & & & & \ddots & \\ & & & a_0 & a_1 & & \cdots & & a_\mu \\ b_0 & b_1 & & \cdots & & b_{n-\mu} & & & \\ & b_0 & b_1 & & \cdots & & b_{n-\mu} & & \\ & & \ddots & \ddots & & & & \ddots & \\ & & & b_0 & b_1 & & \cdots & & b_{n-\mu} \end{bmatrix} \begin{bmatrix} s^{n-1} \\ s^{n-2}t \\ \vdots \\ s^\mu t^{n-\mu-1} \\ \vdots \\ st^{n-2} \\ t^{n-1} \end{bmatrix}$$

Hence the matrix M_{n-1} in Theorem 3.15 is the Sylvester matrix for the resultant we saw in Section 1.1. It follows that

$$\det M_{n-1} = \operatorname{Res}(p, q) = F^{\deg(\Phi)},$$

which proves Theorem 3.15 for the special basis (3.26) of $\operatorname{Syz}(a, b, c)_{n-1}$. The general case follows since changing to an arbitrary basis multiplies $\det M_{n-1}$ by a nonzero constant.

The Rees Algebra. Given an ideal I in a commutative ring R, its *Rees algebra* is the graded R-algebra

$$\mathcal{R}(I) = \bigoplus_{m=1}^{\infty} I^m e^m \subseteq R[e].$$

If $I = \langle f_1, \dots, f_d \rangle$, then $x_i \mapsto f_i e$ induces a surjection

$$R[y_1, \dots, y_d] \xrightarrow{\beta} \mathcal{R}(I) \subseteq R[e].$$

Generators of the kernel $K = \ker(\beta)$ are called *defining equations* of $\mathcal{R}(I)$. When we use the standard grading on $R[y_1, \dots, y_d]$, we see that K is also graded, with K_j denoted the elements of K of degree j in $y_1, \dots, y_d$.

The Rees algebra of an ideal $I = \langle f_1, \dots, f_d \rangle$ has a nice relation to the syzygy module of its generators $f_1, \dots, f_d$:

- $(h_1, \dots, h_d) \in \operatorname{Syz}(f_1, \dots, f_d)$ if and only if $h_1 y_1 + \cdots + h_d y_d \in K_1$. It follows that $\operatorname{Syz}(f_1, \dots, f_d) \simeq K_1$.
- K_1 generates K when $f_1, \dots, f_d$ form a regular sequence in R.
- The quotient ring $R[y_1, \dots, y_d]/\langle K_1 \rangle$ is isomorphic to the *symmetric algebra* $\operatorname{Sym}_R(I)$, which will play a role in the next section.

In what follows, we will focus on the Rees algebra of $I = \langle a, b, c \rangle \subseteq R = k[s, t]$. Using the coordinates x, y, z of $\mathbb{P}^2$, we write the presentation of $\mathcal{R}(I)$ as

$$0 \longrightarrow K \longrightarrow R[x, y, z] \xrightarrow{\beta} \mathcal{R}(I) \longrightarrow 0, \tag{3.27}$$

where β maps x, y, z to $ae, be, ce \in R[e]$.

Moving Curves and the Rees Algebra. The ideal of defining equations K in (3.27) was discovered by the geometric modeling community, as we now explain.

Let $B \in k[s,t;x,y,z]$ be *bihomogeneous* of degree i in s,t and degree j in x,y,z, written $\deg(B) = (i,j)$. Then:

- For a fixed $(s,t) \in \mathbb{P}^1$, the equation $B(s,t,x,y,z) = 0$ defines a curve in $\mathbb{P}^2$ of degree j.
- When we vary $(s,t) \in \mathbb{P}^1$, we get a *moving curve* of degree j in $\mathbb{P}^2$.

In [**314**], Sederberg and Chen consider moving curves that *follow the parametrization*, meaning

$$B(s,t,a(s,t),b(s,t),c(s,t)) \equiv 0.$$

Since B is homogeneous of degree j in x,y,z, this implies

$$0 = e^j B(s,t,a(s,t),b(s,t),c(s,t)) = B(s,t,a(s,t)e,b(s,t)e,c(s,t)e) = \beta(B),$$

where β is from (3.27). Hence $B \in K_j$, which easily leads to the following result.

LEMMA 3.17. *If $B \in R[x,y,z]$ is bihomogeneous of degree (i,j), then the moving curve $B = 0$ follows the parametrization $\Phi = (a,b,c)$ if and only if $B \in K_j$.*

Computing the Defining Equations. In the larger ring $R[e,x,y,z]$, the map

$$R[e,x,y,z] \longrightarrow R[e]$$

sending e,x,y,z to e,ae,be,ce has kernel $\langle x-ae, y-be, z-ce\rangle$. It follows that $K = \ker(\beta : R[x,y,z] \to \mathcal{R}(I) \subseteq R[e])$ is the intersection

$$K = R[x,y,z] \cap \langle x-ae, y-be, z-ce\rangle \subseteq R[e,x,y,z].$$

Since $R = k[s,t]$, this becomes

$$K = k[s,t,x,y,z] \cap \langle x-ae, y-be, z-ce\rangle \subseteq k[e,s,t,x,y,z],$$

which is a problem in elimination theory. Here is an example.

EXAMPLE 3.18. For the parametrization with $n = 4$ given by

$$\Phi = (a,b,c) = (t^4, s^2t^2, s^4 - st^3),$$

a computation in MACAULAY2 gives the minimal generators of K:

```
i1 : R = QQ[e,s,t,x,y,z, MonomialOrder => Eliminate 1]

i2 : J = ideal (x - t^4*e, y - s^2*t^2*e, z - (s^4-s*t^3)*e)

i3 : mingens ideal selectInSubring(1, generators gb J)

o3 = |txy-sy2+sxz sx2-ty2+txz stx-s2y+t2z s2x-t2y x3y-y4+2xy2z-x2z2|
```

moving conics | moving lines | implicit equation

Recall that (i,j) means degree i in s,t and degree j in x,y,z. Then K has five minimal generators with degrees as follows:

- Two moving conics of degree $(1,2)$.
- Two moving lines of degree $(2,1)$. Thus $\mu = 2$.
- The implicit equation is a quartic. Degree $(0,4)$ means that it is the moving curve that doesn't move!

We will soon see that these degrees are no accident when $n = 4$ and $\mu = 2$. The next section will present a slicker way of computing K in MACAULAY2. ◁▷

Two Theorems about Minimal Generators. The first theorem is from [**313**], where Sederberg, Goldman and Du generalize Example 3.18 as follows.

THEOREM 3.19. *If a curve parametrization* $\Phi : \mathbb{P}^1 \xrightarrow{(a,b,c)} \mathbb{P}^2$ *has* $n = 4$ *and* $\deg(\Phi) = 1$*, then:*

(1) $\mu = 2$ *if and only if all singular points of* C *have multiplicity* 2.
(2) *When* $\mu = 2$*, the defining equations ideal* K *has five minimal generators:*
   – *Generators* p *and* q *of degree* $(2,1)$ *(the* μ*-basis).*
   – *Two moving conics* C_1 *and* C_2 *of degree* $(1,2)$*.*
   – *The implicit equation* $F = 0$ *of degree* $(0,4)$*.*

A rigorous proof using local duality can be found in [**96**]. This paper also reveals the following "internal structure" of the minimal generators of K. The first step is to write the μ-basis in terms of s^2, st, t^2:

$$p = s^2\ell_0(x,y,z) + st\ell_1(x,y,z) + t^2\ell_2(x,y,z)$$
$$q = s^2\ell'_0(x,y,z) + st\ell'_1(x,y,z) + t^2\ell'_2(x,y,z),$$

where $\ell_0, \ldots, \ell'_2$ are linear in x, y, z. Then compute polynomials C_1, C_2, F as follows:

- Express p, q in terms of s^2, t to obtain

$$\begin{matrix} p = s^2\ell_0 + t(s\ell_1 + t\ell_2) \\ q = s^2\ell'_0 + t(s\ell'_1 + t\ell'_2) \end{matrix} \text{ and define } C_1 = \det\begin{bmatrix} \ell_0 & s\ell_1 + t\ell_2 \\ \ell'_0 & s\ell'_1 + t\ell'_2 \end{bmatrix}.$$

  Note that C_1 has degree $(1,2)$.
- Express p, q in terms of s, t^2 to obtain

$$\begin{matrix} p = s(s\ell_0 + t\ell_1) + t^2\ell_2) \\ q = s(s\ell'_0 + t\ell'_1) + t^2\ell'_2) \end{matrix} \text{ and define } C_2 = \det\begin{bmatrix} s\ell_0 + t\ell_1 & \ell_2 \\ s\ell'_0 + t\ell'_1 & \ell'_2 \end{bmatrix}.$$

  Similarly, C_2 has degree $(1,2)$.
- Express C_1, C_2 in terms of s, t to obtain

$$\begin{matrix} C_1 = sQ_0(x,y,z) + tQ_1(x,y,z) \\ C_2 = sQ'_0(x,y,z) + tQ'_1(x,y,z) \end{matrix} \text{ and define } F = \det\begin{bmatrix} Q_0 & Q_1 \\ Q'_0 & Q'_1 \end{bmatrix}.$$

  Note that F has degree $(0,4)$.

Theorem 2.12 of [**96**] shows that in Theorem 3.19, minimal generators of K are given by the μ-basis p, q together with C_1, C_2, F as defined above. The last three polynomials are examples of *Sylvester forms*. Thus one can construct the defining relation ideal K, including the implicit equation, from the μ-basis via iterated Sylvester forms.

The second theorem shows that the case $\mu = 1$ follows a similar pattern.

THEOREM 3.20. *If a curve parametrization* $\Phi : \mathbb{P}^1 \xrightarrow{(a,b,c)} \mathbb{P}^2$ *has* $n \geq 3$ *and* $\deg(\Phi) = 1$*, then:*

(1) $\mu = 1$ *if and only if* C *has a unique singular point of multiplicity* $n - 1$*.*
(2) *When* $\mu = 1$*, the defining equations ideal* K *has* $n + 1$ *minimal generators consisting of the* μ*-basis elements* p, q *of degrees* $(1,1)$ *and* $(n-1,1)$*, together with polynomials* q_i *of degree* $(n-i, i)$ *for* $i = 2, \ldots, n$*. Furthermore:*
   – $F = q_n$ *has degree* $(0,n)$ *and gives the implicit equation of* C*.*
   – *If we set* $q_1 = q$*, then for* $i = 2, \ldots, n$*,* q_i *can be obtained as the Sylvester form of* p *and* q_{i-1}*.*

This result was conjectured in 2007 by Hong, Simis and Vasconcelos [**210**] and proved by Cox, Hoffman and Wang [**101**] a year later.

Many people have worked in the area, including Busé, Cortadellas, D'Andrea, Jia, Kustin, Madsen, Polini, Ulrich and others. See the paper [**92**] for recent results and references to the literature. There are also lovely relations to singularities and interesting generalizations to parametrizations $\mathbb{P}^1 \to \mathbb{P}^N$. The paper [**102**] has numerous useful citations, including the pioneering papers of Maria-Grazia Ascenzi from the late 1980s.

Surface Parametrizations. It should be no surprise that the algebra involved in surface parametrizations is richer and more complicated than for curves. We begin our discussion of surfaces with an example that illustrates some of the new features to consider.

EXAMPLE 3.21. Let $\mathbb{A}^2$ denote 2-dimensional affine space, typically $\mathbb{R}^2$ or $\mathbb{C}^2$, and similarly for $\mathbb{A}^3$. Then consider the parametrization

$$(s,t) \in \mathbb{A}^2 \longmapsto (x,y,z) = \Big(\frac{s}{s^2+t^2}, \frac{t}{s^2+t^2}, \frac{st}{s^2+t^2}\Big) \in \mathbb{A}^3$$

For curves, the link to algebra worked best in the projective context. But what does "projective" mean in the surface case? Here are two ways to make the above map projective:

$$\begin{aligned} (s,t,u) \in \mathbb{P}^2 &\longmapsto (su, tu, st, s^2+t^2) \in \mathbb{P}^3 \\ (s,u;t,v) \in \mathbb{P}^1 \times \mathbb{P}^1 &\longmapsto (suv^2, u^2tv, sutv, s^2v^2+u^2t^2) \in \mathbb{P}^3. \end{aligned} \tag{3.28}$$

In the language of Section 3.1, the first is a triangular parametrization of degree 2, while the second is a tensor product parametrization of degree $(2,2)$.

Notice also the presence of *basepoints* in (3.28), which are parameter values where the parametization is undefined as a point in $\mathbb{P}^3$. One can compute that

- $\mathbb{P}^2 \dashrightarrow \mathbb{P}^3$ has basepoint $(0,0,1)$.
- $\mathbb{P}^1 \times \mathbb{P}^1 \dashrightarrow \mathbb{P}^3$ has basepoints $(0,1;0,1)$ and $(1,0;1,0)$.

Even for a parametrization $\mathbb{R}^2 \to \mathbb{R}^3$ (typical in geometric modeling), we will see that basepoints at ∞ or over $\mathbb{C}$ cannot be ignored. ◁▷

In the affine setting, a surface parametrization is

$$\Phi : \mathbb{A}^2 \dashrightarrow \mathbb{A}^3, \quad \Phi(s,t) = \Big(\frac{a(s,t)}{d(s,t)}, \frac{b(s,t)}{d(s,t)}, \frac{c(s,t)}{d(s,t)}\Big),$$

where $a,b,c,d \in k[s,t]$ are relatively prime and $k = \mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$. We will often write this as

$$\Phi : \mathbb{A}^2 \dashrightarrow \mathbb{P}^3, \quad \Phi(s,t) = \big(a(s,t), b(s,t), c(s,t), d(s,t)\big),$$

denoted $\Phi = (a,b,c,d)$. We always assume that the Zariski closure of the image of Φ is a surface $S \subseteq \mathbb{P}^3$. For simplicity, we focus on two ways to make this projective:

- Triangular: $\Phi : \mathbb{P}^2 \dashrightarrow \mathbb{P}^3$ is given by $a,b,c,d \in k[s,t,u]$, homogeneous of the same degree n.
- Tensor product: $\Phi : \mathbb{P}^1 \times \mathbb{P}^1 \dashrightarrow \mathbb{P}^3$ is given by $a,b,c,d \in k[s,u;t,v]$, bihomogeneous of the same degree (m,n).

In each case, we get an ideal I and a basepoint locus $\mathcal{B}$:

- Triangular: $I = \langle a, b, c, d\rangle \subseteq k[s, t, u]$ with $\mathcal{B} = \mathbf{V}(I) \subseteq \mathbb{P}^2$.
- Tensor product: $I = \langle a, b, c, d\rangle \subseteq k[s, u; t, v]$ with $\mathcal{B} = \mathbf{V}(I) \subseteq \mathbb{P}^1 \times \mathbb{P}^1$.

For curves, the assumption $\gcd(a, b, c) = 1$ implies that there are no basepoints, while for surfaces, $\gcd(a, b, c, d) = 1$ guarantees that the basepoint locus $\mathcal{B}$ is finite.

The Degree Formula. Similar to the curve case, the degree formula for a surface parametrization $\Phi = (a, b, c, d)$ involves three types of degree:

- a, b, c, d have degree n (triangular) or degree (m, n) (tensor product).
- $\deg(S)$ is the degree of the irreducible homogeneous $F \in k[x, y, z, w]$ with $S = \mathbf{V}(F)$.
- $\deg(\Phi)$ is the number of points in the domain that map to a generic point of S when working over $\mathbb{C}$.

In addition, each basepoint $p \in \mathcal{B} = \mathbf{V}(I)$ contributes to the degree formula via its *Hilbert-Samuel multiplicity* $e_p(I)$. We refer the reader to [**139**, Chapter 12] for the definition of $e_p(I)$.

We are now ready to state the degree formula for a surface parametrization.

THEOREM 3.22. *If $a, b, c, d \in k[s, t, u]$ are relatively prime and homogeneous of degree n, then the parametrization $\Phi : \mathbb{P}^2 \dashrightarrow S \subseteq \mathbb{P}^3$ satisfies*

$$\deg(\Phi)\deg(S) = n^2 - \sum_{p\in\mathcal{B}} e_p(I).$$

On the other hand, if $a, b, c, d \in k[s, u; t, v]$ are relatively prime and bihomogeneous of degree (m, n), then $\Phi : \mathbb{P}^1 \times \mathbb{P}^1 \dashrightarrow S \subseteq \mathbb{P}^3$ satisfies

$$\deg(\Phi)\deg(S) = 2mn - \sum_{p\in\mathcal{B}} e_p(I).$$

A proof for triangular parametrizations can be found in [**94**]. This proof is easily adapted to tensor product case.

EXAMPLE 3.23. For the affine parametrization

$$x = \frac{s}{s^2 + t^2}, \quad y = \frac{t}{s^2 + t^2}, \quad z = \frac{st}{s^2 + t^2}$$

from Example 3.21, elimination theory gives $S \subseteq \mathbb{A}^3$ defined by $x^2z+y^2z-2xy = 0$. Hence $\deg(S) = 3$. Let's compute this using the degree formula:

- $\Phi : \mathbb{P}^2 \dashrightarrow S \subseteq \mathbb{P}^3$ defined by $\Phi(s, t, u) = (su, tu, st, s^2+t^2)$ has degree 2 and $\deg(\Phi) = 1$. The basepoint $p = (0, 0, 1)$ has $e_p(I) = 1$, so that
$$\deg(S) = 2^2 - 1 = 3.$$
- $\Phi : \mathbb{P}^1\times\mathbb{P}^1 \dashrightarrow S \subseteq \mathbb{P}^3$ defined by $\Phi(s, u; t, v) = (suv^2, u^2tv, sutv, s^2v^2+u^2t^2)$ has degree $(2, 2)$ and $\deg(\Phi) = 1$. As for the basepoints, $p_1 = (0, 1; 0, 1)$ has $e_{p_1}(I) = 1$ and $p_2 = (1, 0; 1, 0)$ has $e_{p_2}(I) = 4$, so that
$$\deg(S) = 2 \cdot 2 \cdot 2 - 1 - 4 = 3.$$

These two computations of $\deg(S)$ illustrate how the basepoints of Φ can differ in number and complexity depending on how Φ is represented. ◁▷

EXAMPLE 3.24. Consider a bicubic parametrization $\Phi = (a, b, c, d)$, such as the one pictured in (3.7). When a, b, c, d are generic of degree $(3, 3)$, one can show that $\mathcal{B} = \mathbf{V}(a, b, c, d) = \emptyset$ and $\deg(\Phi) = 1$, so the parametrized surface S has degree

$$\deg(S) = 2 \cdot 3 \cdot 3 = 18.$$

Recall from Section 1.2 that geometric modelers first computed this number via ray tracing. It is nice to see how algebraic geometry gives the same answer.

It follows that the implicit equation $F = 0$ of S will typically have $\binom{21}{18} = 1330$ terms. This is small by modern standards but not something one wants to write out explicitly. ◁▷

Syzygies and Moving Planes. Syzygies, moving lines, and μ-bases are useful for curves. This is true for surfaces, with some complications. Let $\Phi : \mathbb{A}^2 \dashrightarrow \mathbb{P}^3$ be given by relatively prime polynomials $a, b, c, d \in k[s, t]$. Then a moving plane in $\mathbb{P}^3$

$$Ax + By + Cz + Dw = 0, \quad A, B, C, D \in k[s, t],$$

follows Φ if and only if $Aa + Bb + Cc + Dd = 0$. Thus moving planes correspond to elements of the syzygy module

$$\mathrm{Syz}(a, b, c, d) \subseteq k[s, t]^4.$$

A first complication is that when we work homogeneously, the syzygy module in $k[s, t, u]^4$ or $k[s, u; t, v]^4$ is rarely free (see, for example, [**94**, Theorem 5.1]). But it is always free in the affine setting, as proved in [**76**].

THEOREM 3.25. *For $a, b, c, d \in k[s, t]$ with $\gcd(a, b, c, d) = 1$, the syzygy module $\mathrm{Syz}(a, b, c, d)$ is a free $k[s, t]$-module of rank 3.*

A basis $\{p, q, r\}$ of $\mathrm{Syz}(a, b, c, d) \subseteq k[s, t]^4$ is called an *affine μ-basis*, which we write either as the matrix

$$\begin{bmatrix} p \\ q \\ r \end{bmatrix} = \begin{bmatrix} A & B & C & D \\ A' & B' & C' & D' \\ A'' & B'' & C'' & D'' \end{bmatrix}. \tag{3.29}$$

or as linear forms

$$\begin{aligned} p &= A\ x + B\ y + C\ z + D\ w \\ q &= A'\, x + B'\, y + C'\, z + D'\, w \\ r &= A''x + B''y + C''z + D''w \end{aligned} \tag{3.30}$$

with coefficients $A, B, \ldots, C'', D'' \in k[s, t]$. Here is an example.

EXAMPLE 3.26. $\Phi(s, t) = (2st, 2t, 2s, s^2 + t^2 + 1)$ has affine μ-basis

$$\begin{aligned} p &= x - sy \\ q &= (t^2 + 1)y + stz - 2tw \\ r &= sty + (s^2 + 1)z - 2sw. \end{aligned} \tag{3.31}$$

As a map $\mathbb{P}^2 \to \mathbb{P}^3$ (no basepoints in this case), the parametrization is

$$\Phi(s, t, u) = (2st, 2tu, 2su, s^2 + t^2 + u^2).$$

However, the homogeneous syzygy module in $k[s, t, u]^4$ is *not free*—one can check using MACAULAY2 that it has five minimal generators! ◁▷

In the affine situation, we also have the Hilbert-Burch Theorem.

THEOREM 3.27. *The* 3×3 *minors of* (3.29) *are* a, b, c, d *up to a constant.*

The Resultant of a μ-Basis. For curves, Theorem 3.16 tells us that the resultant of a μ-basis is the implicit equation when $\deg(\Phi) = 1$. We now explore the extent to which this remains true in the surface case. Given an affine μ-basis $\{p, q, r\} \subseteq k[s, t]^4$, there are three related issues to consider.

ISSUE 1: Resultants require homogeneous inputs. Hence:

- For a triangular parametrization $\Phi : \mathbb{P}^2 \to \mathbb{P}^3$, make the polynomials p, q, r in (3.30) homogeneous with respect to s, t, u. The resultant is
$$\operatorname{Res}(p, q, r) = \operatorname{Res}_{s,t,u}(p(s, t, u), q(s, t, u), r(s, t, u)),$$
where $\operatorname{Res}_{s,t,u}$ is the classical resultant from (1.11) in Section 1.2.
- For a tensor product parametrization $\Phi : \mathbb{P}^1 \times \mathbb{P}^1 \to \mathbb{P}^3$, make p, q, r bihomogeneous with respect to s, u and t, v. The resultant is
$$\operatorname{Res}(p, q, r) = \operatorname{Res}_{s,u;t,v}(p(s, u; t, v), q(s, u; t, v), r(s, u; t, v)),$$
where $\operatorname{Res}_{s,u;t,v}$ is the sparse or Dixon resultant discussed in Section 1.2.

$\operatorname{Res}(p, q, r)$ is homogeneous in x, y, z, w and vanishes on S as we will soon see.

ISSUE 2: Basepoints have a role to play. By Hilbert-Burch, $a, b, c, d \in k[s, t]$ are the maximal minors of the 3×4 matrix (3.29) with rows p, q, r. Hence

$$\begin{aligned}(s_0, t_0) \in \mathbb{A}^2 \text{ is a basepoint } &\iff \text{(3.29) has rank} < 3 \text{ at } (s_0, t_0)\\ &\iff \text{(3.29) has rank} = 0, 1, 2 \text{ at } (s_0, t_0).\end{aligned} \tag{3.32}$$

Basepoints can lead to *extraneous factors* in $\operatorname{Res}(p, q, r)$ as we will see below.

ISSUE 3: As the name implies, an affine μ-basis $\{p, q, r\}$ gives no information about how the corresponding moving planes behave at ∞ (meaning $\mathbb{P}^2 \setminus \mathbb{A}^2$ or $\mathbb{P}^1 \times \mathbb{P}^1 \setminus \mathbb{A}^2$). There are two things that can happen at ∞:

- Basepoints can occur at ∞, sometimes leading to extraneous factors.
- Failure to generate the syzygy module at ∞ can lead to extraneous factors.

EXAMPLE 3.28. Written in homogeneous coordinates, Example 3.26 gives the parametrization
$$\Phi(s, t, u) = (2st, 2tu, 2su, s^2 + t^2 + u^2).$$
As previously noted, there are no basepoints. Yet the μ-basis (3.31) has resultant
$$\operatorname{Res}(p, q, r) = y^4 F(x, y, z, w),$$
where $F = 0$ is the implicit equation of the surface $S \subseteq \mathbb{P}^3$. The extraneous factor y^4 arises from the failure of (3.31) to be a μ-basis at ∞. ◁▷

Here is an example where the resultant is useless.

EXAMPLE 3.29. The affine parametrization
$$x = \frac{s^3 + t^4}{t^3}, \quad y = \frac{s^2 t}{t^3} = \frac{s^2}{t^2}, \quad z = \frac{st^2}{t^3} = \frac{s}{t}$$
has $\operatorname{Res}(p, q, r) \equiv 0$. We will see in Example 3.31 below that this happens because $(s, t) = (0, 0)$ is a *really bad* basepoint. ◁▷

A First Look at Res(p, q, r). In spite of the problems just noted, Res(p, q, r) has some nice properties. To understand why, observe that Res(p, q, r) vanishes at any point (x, y, z, w) where the equations

$$\begin{aligned} p &= A\ x + B\ y + C\ z + D\ w = 0 \\ q &= A'\, x + B'\, y + C'\, z + D'\, w = 0 \\ r &= A''x + B''y + C''z + D''w = 0 \end{aligned} \tag{3.33}$$

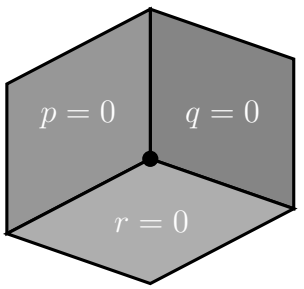
have a solution s, t. Since $(x, y, z, w) = \Phi(s, t)$ is a solution, we have the following.

THEOREM 3.30. Res(p, q, r) *vanishes on the parametrized surface* $S \subseteq \mathbb{P}^3$.

Furthermore, if $(s, t) \in \mathbb{A}^2$ is not at a basepoint, then the matrix

$$\begin{bmatrix} p \\ q \\ r \end{bmatrix} = \begin{bmatrix} A & B & C & D \\ A' & B' & C' & D' \\ A'' & B'' & C'' & D'' \end{bmatrix}$$

has rank 3 at (s, t) since a, b, c, d are its 3×3 minors by Hilbert-Burch. This shows that away from the basepoints, the moving planes have a unique point of intersection:



This point must be on S since the planes follow the parametrization. Thus we recover the surface analog of the picture for curves shown in (3.22).

However, Examples 3.26 and 3.29 show that Res(p, q, r) may have extraneous factors or vanish identically. We now give an informal discussion of how basepoints affect the resultant.

Behavior at Basepoints. At a basepoint, the parameter values "blow up" to an *exceptional curve* on the surface $S \subseteq \mathbb{P}^3$. Exceptional curves come in three flavors:

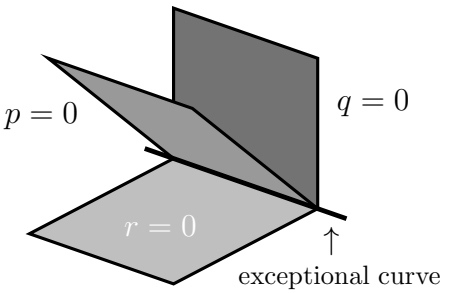
- A line.
- A plane curve.
- A space curve.

On the other hand, at a basepoint, we noted earlier that the matrix

$$\begin{bmatrix} p \\ q \\ r \end{bmatrix} = \begin{bmatrix} A & B & C & D \\ A' & B' & C' & D' \\ A'' & B'' & C'' & D'' \end{bmatrix}$$

has rank < 3, so the rank is 2, 1 or 0. A wonderful fact is that the three flavors of exceptional curves correspond to rank $= 2, 1, 0$. Let's explore each of these, starting with the nicest.

At a Rank 2 Basepoint. Here, the moving planes intersect in a line:



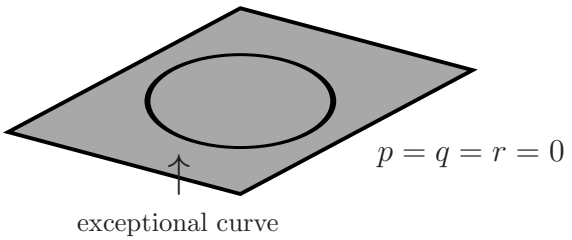
The line lies on the surface and is the exceptional curve. There is no extraneous factor to worry about.

We also have a nice connection to commutative algebra:

$$p = (s,t) \text{ is a rank 2 basepoint} \iff \begin{array}{l}\text{the localized ideal } \langle a,b,c,d\rangle \subseteq \mathcal{O}_{\mathbb{A}^2,p} \\ \text{has two minimal generators.}\end{array}$$

In other words, a basepoint has rank 2 if and only if the the ideal $\langle a,b,c,d\rangle$ is a *local complete intersection* (LCI) at the basepoint.

At a Rank 1 Basepoint. Here, the three moving planes coincide:



The exceptional curve is a curve lying in this plane. The resultant acquires an extraneous factor that is the equation of the plane raised to some exponent.

To explain the exponent, we need more commutative algebra. At a basepoint $p = (s,t)$ of the ideal $I = \langle a,b,c,d\rangle$, we have two numerical invariants:

- The *Hilbert-Samuel multiplicity* $e_p(I)$ that appears in the degree formula.
- The *degree* $d_p(I) = \dim \mathcal{O}_{\mathbb{A}^2,p}/\langle a,b,c,d\rangle$.

These invariants satisfy

$$e_p(I) \geq d_p(I),$$

with equality if and only if p is LCI. We also know that

$$p = (s,t) \text{ is a rank 1 basepoint} \iff \begin{array}{l}\text{the localized ideal } \langle a,b,c,d\rangle \subseteq \mathcal{O}_{\mathbb{A}^2,p} \\ \text{has three minimal generators.}\end{array}$$

This is called a *local almost complete intersection*. At a rank 1 basepoint p, the extraneous factor of $\mathrm{Res}(p,q,r)$ is given by

$$L_p^{e_p(I)-d_p(I)},$$

where $L_p = 0$ is the plane defined by the μ-basis at the basepoint. The exponent $e_p(I) - d_p(I) > 0$ is a numerical measure of how close a local almost complete intersection is to being LCI.

At a Rank 0 Basepoint. Here, the moving "planes" are the ambient space since *every* coefficient vanishes. In this case, the exceptional curve is a space curve. At such a basepoint, the equations (3.33) have a solution for *any* (x, y, z, w), so that the resultant $\operatorname{Res}(p, q, r)$ vanishes identically. Finally,

$$p = (s, t) \text{ is a rank 0 basepoint} \iff \begin{array}{l}\text{the localized ideal } \langle a, b, c, d\rangle \subseteq \mathcal{O}_{\mathbb{A}^2,p} \\ \text{has four minimal generators.}\end{array}$$

Rank 0 basepoints have no special name beyond the fact that they are really bad.

EXAMPLE 3.31. The parametrization

$$x = \frac{s^3 + t^4}{t^3}, \quad y = \frac{s^2 t}{t^3} = \frac{s^2}{t^2}, \quad z = \frac{st^2}{t^3} = \frac{s}{t}$$

from Example 3.29 has a basepoint at $(s, t) = (0, 0)$. Here, the ideal is

$$\langle s^3 + t^4, s^2 t, st^2, t^3\rangle = \langle s^3, s^2 t, st^2, t^3\rangle = \langle s, t\rangle^3,$$

which requires four minimal generators, even after localization. This shows that $(0, 0)$ is a rank 0 basepoint and explains why $\operatorname{Res}(p, q, r) \equiv 0$. ◁▷

What We Have and What We Want. Summarizing the above discussion, we see that if all basepoints are at worst local almost complete intersections, then the resultant of an affine μ-basis can be written

$$\operatorname{Res}(p, q, r) = \underbrace{F^{\deg(\Phi)} \times \prod_{\text{rank 1}} L_p^{e_p(I)-d_p(I)}}_{\text{what we want}} \times \underbrace{\text{extraneous factor at } \infty}_{\text{known by [58] and [319]}},$$

where the product is over all rank 1 basepoints and $F = 0$ is the implicit equation of the surface S. The theory behind this formula is explained by Busé, Chardin and Jouanolou [**58**], who also analyze the extraneous factors at ∞ in the triangular case. Extraneous factors at ∞ in the tensor product case are described by Shen and Goldman [**319**]. Their paper also includes some lovely examples.

The drawback of this approach is that μ-bases and resultants are not easy to compute, and there are often extraneous factors at ∞ we don't want. As indicated above, what we really want is the product

$$F^{\deg(\Phi)} \times \prod_{\text{rank 1}} L_p^{e_p(I)-d_p(I)}. \tag{3.34}$$

We next give an alternate approach to this formula based on *matrix representations* (also called *implicitization matrices*) built from moving planes of fixed degree.

Matrix Representations. We begin with a definition.

DEFINITION 3.32. A homogeneous polynomial $H \in k[x, y, z, w]$ is *represented by a matrix* M of homogeneous polynomials in $k[x, y, z, w]$ when:

(1) The rank of the matrix M drops precisely when evaluated at the points of the surface in $\mathbb{P}^3$ (over $\overline{k}$) defined by $H = 0$.
(2) H is the gcd of the maximal minors of M.

In [**58**], Busé, Chardin and Jouanolou explain how to construct a matrix that represents (3.34) in the triangular case. Let $a, b, c, d \in k[s, t, u]$ be relatively prime and homogeneous degree of degree n. For an integer ν, let $S_1, \ldots, S_N$ be a basis of $\operatorname{Syz}(a, b, c, d)_\nu \subseteq k[s, t, u]^4_\nu$. This gives N moving planes

$$\mathcal{L}_i = S_i \cdot (x, y, z, w), \quad i = 1, \ldots, N.$$

As we have done several times before, express $\mathcal{L}_i$ as a polynomial in s, t, u whose coefficients are linear forms in x, y, z, w. These linear forms give a matrix M_ν with the property that

$$M_\nu \begin{bmatrix} \vdots \\ s^\alpha t^\beta u^\gamma \\ \vdots \end{bmatrix} = \begin{bmatrix} \mathcal{L}_1 \\ \vdots \\ \mathcal{L}_N \end{bmatrix},$$

where the column vector on the left consists of all monomials of degree ν in s, t, u. The following result from [**58**] shows that M_ν is the matrix we want.

THEOREM 3.33. *Assume that k algebraically closed and that the basepoints of $\Phi = (a, b, c, d)$ are at worst local almost complete intersections. If $\nu \geq 2(n-1)$, then the above matrix M_ν represents the polynomial $F^{\deg(\Phi)} \times \prod_{\text{rank } 1} L_p^{e_p(I)-d_p(I)}$.*

This result dates from 2009, though the idea of dropping rank goes back to van der Waerden's 1926 proof of the Fundamental Theorem of Elimination Theory [**351**]. Here are two further comments about the paper [**58**]:

- The proof of Theorem 3.33 in [**58**] uses the *approximation complexes* of Herzog, Simis and Vasconcelos [**202**], which were first applied to implicitization in 2003 by Busé and Jouanolou [**64**]. See Section 3.3 for more details.
- The result in [**58**] generalizes to parametrizations $\mathbb{P}^m \dashrightarrow \mathbb{P}^{m+1}$. The curve case is $m = 1$, where the matrix M_ν works for $\nu \geq n - 1$. If $\nu = n - 1$, then $M_\nu = M_{n-1}$ is the square matrix that computes $\operatorname{Res}(p, q)$ in Theorem 3.15.

Applications of Matrix Representations. In 2009, Luu Ba, Busé and Mourrain [**263**] used matrix representations and numerical linear algebra to study intersections of a curve and a surface. This was extended to the intersection of two surfaces by Busé and Luu Ba in [**56**].

There is also a version for $\mathbb{P}^1 \times \mathbb{P}^1$ or more generally for $\Phi : X_P \dashrightarrow \mathbb{P}^3$ where X_P is the toric surface of the Newton polygon P of a, b, c, d. This brings us back to a result mentioned in the *Preview of the 21st Century* at the end of Section 1.2. In the notation of this section, the goal is to eliminate s, t from

$$x = \frac{a(s,t)}{d(s,t)}, \quad y = \frac{b(s,t)}{d(s,t)}, \quad z = \frac{c(s,t)}{d(s,t)}. \tag{3.35}$$

Pick a lattice polygon $P \subseteq \mathbb{Z}^2_{\geq 0}$ that contains the Newton polytopes of a, b, c, d. If X_P is the toric surface of P, then a, b, c, d give a parametrization

$$\Phi : X_P \dashrightarrow \mathbb{P}^3.$$

To find the implicit equation $F = 0$ of the image, we proceed as follows. The set of all syzygies of $Aa + Bb + Cc + Dd = 0$ where the Newton polytopes of A, B, C, D are contained in $2P$ is a finite-dimensional vector space that can be computed by linear algebra. A basis of such syzygies gives polynomials

$$\mathcal{L}^{(l)} = A^{(l)}(s,t)x + B^{(l)}(s,t)y + C^{(l)}(s,t)z + D^{(l)}(s,t)w, \quad l = 1, \ldots, N,$$

which we rewrite in the form

$$\mathcal{L}^{(l)} = \sum_{(i,j)\in 2P\cap\mathbb{Z}^2} \ell^{(l)}_{(i,j)}(x,y,z,w)\, s^i t^j, \quad l = 1,\dots,N,$$

for linear forms $\ell^{(l)}_{(i,j)}(x,y,z,w)$. When we arrange the $\ell^{(l)}_{(i,j)}$ into a matrix M of linear forms, we have the following result of Botbol and Dickenstein [**41**].

THEOREM 3.34. *If all basepoints in X_P are LCI, then the matrix M represents $F^{\deg(\Phi)}$, where the implicit equation $F = 0$ of the surface S parametrized by* (3.35).

This is Theorem 1.18 from Section 1.2, though in the earlier version, we had to say "in nice cases" rather than "LCI basepoints" and say "drops rank precisely when evaluated at points on the surface" rather than "represents $F^{\deg(\Phi)}$."

Final Comments. The algebra of geometric modeling has a very different flavor from the geometry discussed in Section 3.1. We met some significant new players, including basepoints, multiplicities, syzygies, and Rees algebras, along with links to earlier topics such as elimination theory, resultants, and toric varieties.

As you will learn in the next section, the algebra keeps getting better.

3.3. Rees Algebras, Syzygies, and Computational Geometry

by Hal Schenck

Our goal in this section is to describe in more depth the ideas from earlier in the chapter and introduce the concepts needed to unleash their full power:

- Rees and Symmetric Algebras.
- Local Cohomology and the Jacobian Dual.
- Fitting ideals and the Annihilator of a Module.
- Free resolutions and MacRae Invariant.
- The Approximation Complex.
- Multigraded Implicitization.
- Extensions and Future Directions.

The setup is that the ideal

$$I = \langle f_1,\dots,f_d\rangle \subseteq R = k[x_1,\dots,x_m]$$

is generated by homogeneous elements of degree n, which define a map of rings

$$S = k[y_1,\dots,y_d] \xrightarrow{\phi} R,\ y_i \longmapsto f_i. \tag{3.36}$$

This in turn gives rise to a corresponding map of varieties

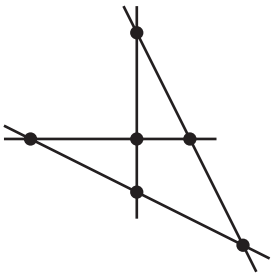
$$\mathbb{P}^{m-1} \setminus \mathcal{B} \xrightarrow{\Phi} \mathbb{P}^{d-1},$$

where the *base locus* $\mathcal{B} = \mathbf{V}(I)$ consists of the points where Φ is undefined. Our focus will be on the case where $d = m+1$ and Φ generically finite onto its image, so the image of Φ is a hypersurface $H = \mathbf{V}(F)$. We begin with a motivating example.

EXAMPLE 3.35. Let $l = x_1 + x_2 + x_3$ and

$$I = \langle x_1x_2x_3, x_1x_2l, x_1x_3l, x_2x_3l\rangle \subseteq R = k[x_1,x_2,x_3].$$

The four lines $\mathbf{V}(x_1), \mathbf{V}(x_2), \mathbf{V}(x_3), \mathbf{V}(l)$ depicted here



determine the base locus $\mathcal{B}$, which consists of the six intersection points

$$\mathcal{B} = \{(0,1,0), (0,0,1), (0,1,-1), (1,0,-1), (1,-1,0), (1,0,0)\}.$$

The ideal I has a presentation given by

$$0 \to R(-4)^3 \xrightarrow{\begin{bmatrix} l & 0 & 0 \\ -x_3 & x_3 & 0 \\ 0 & -x_2 & x_2 \\ 0 & 0 & -x_1 \end{bmatrix}} R(-3)^4 \xrightarrow{\begin{bmatrix} x_1x_2x_3 & x_1x_2l & x_1x_3l & x_2x_3l \end{bmatrix}} I \to 0.$$

The leftmost matrix consists of the *syzygies* on the ideal I: polynomial relations on the 1×4 matrix of polynomials generating I. If we denote the basis vectors for R^4 as $\{y_1, \dots y_4\}$ then we may rewrite the syzygies as

$$ly_1 - x_3y_2, x_3y_2 - x_2y_3, x_2y_3 - x_1y_4.$$

Now we turn this inside out, and write the three expressions above (recall that $l = x_1 + x_2 + x_3$) as column vectors with respect to the basis $\{x_1, x_2, x_3\}$ of R_1:

$$B = \begin{bmatrix} y_1 & 0 & -y_4 \\ y_1 & -y_3 & y_3 \\ y_1 - y_2 & y_2 & 0 \end{bmatrix}.$$

Let

$$F = \det(B) = y_2y_3y_4 - y_1y_3y_4 - y_1y_2y_4 - y_1y_2y_3.$$

An easy computation shows that $F(f_1, f_2, f_3, f_4) = 0$, so that the determinant of B is the implicit equation for the hypersurface H.

It is instructive to examine the geometry of this example more closely. It is a classical fact [**190**, §V.3] that the blow up X of $\mathbb{P}^2$ at six general points may be embedded as a smooth cubic surface in $\mathbb{P}^3$, using the divisor

$$D = 3E_0 - \sum_{i=1}^{6} E_i,$$

where E_0 is the proper transform of a line and E_i is the exceptional curve over a blown up point. The intersection pairing on X satisfies

$$E_0^2 = 1 \quad E_i^2 = -1 \text{ if } i \neq 0 \quad E_iE_j = 0 \text{ if } i \neq j.$$

In our example, the points are not general: for example, the line $L = \mathbf{V}(x_1)$ contains the 3 basepoints (say p_1, p_2, p_3), and we compute

$$D \cdot L = (3E_0 - \sum_{i=1}^{6} E_i) \cdot (E_0 - E_1 - E_2 - E_3) = 3 - 3 = 0.$$

This computation shows that any of the four lines which contain three points of the base locus is contracted to a point; in particular we expect $\mathbf{V}(F)$ to have four singular points. The Jacobian ideal of F is generated by

$$\begin{bmatrix} y_2y_3 + y_2y_4 + y_3y_4 \\ y_1y_3 + y_1y_4 - y_3y_4 \\ y_1y_2 + y_1y_4 - y_2y_4 \\ y_1y_2 + y_1y_3 - y_2y_3 \end{bmatrix},$$

which has zeroes at exactly the four coordinate points of $\mathbb{P}^3$, confirming our expectation that $H = \mathbf{V}(F)$ has four singular points. ◁▷

The remainder of this section will be devoted to understanding the tools and techniques that are used to produce the implicit equation F, in particular to explain the mysterious switch between the matrix of syzygies of I, and the matrix B.

Rees and Symmetric Algebras. In the previous section, we defined the Rees algebra $\mathcal{R}(I)$ of an ideal $I = \langle f_1, \ldots, f_d \rangle \subseteq R$; it is the image of the map

$$R[y_1, \ldots, y_d] \xrightarrow{\beta} R[e] \text{ via } y_i \mapsto f_i \cdot e.$$

If R is a domain, then since $\mathcal{R}(I)$ is isomorphic to a subring of $R[e]$, then $\mathcal{R}(I)$ will also be a domain. If we grade by the variables y_i and let $K = \ker(\beta)$, then:

$$\begin{array}{lcll} K_1 & = & \sum a_i y_i \mapsto e \cdot \sum a_i f_i = 0 & \iff (a_i) \in \operatorname{Syz}(I) \\ K_2 & = & \sum a_{ij} y_i y_j \mapsto e^2 \cdot \sum a_{ij} f_i f_j = 0 & \iff (a_{ij}) \in \operatorname{Syz}(I^2) \\ K_3 & = & \vdots & \iff (a_{ijk}) \in \operatorname{Syz}(I^3) \end{array}$$

It follows from the definition that $\mathcal{R}(I) \simeq R[y_1, \ldots, y_d]/K$. A close cousin of $\mathcal{R}(I)$ is the *symmetric algebra* $\operatorname{Sym}_R(I)$, defined as $R[y_1, \ldots, y_d]/\langle K_1 \rangle$. It turns out that computing $\mathcal{R}(I)$ is a difficult problem, discussed at length at the end of the chapter. The symmetric algebra $\operatorname{Sym}_R(I)$ is simpler; in the special case where $K = \langle \ker(\beta)_1 \rangle$, the ideal I is said to be of *linear type*.

Example 3.36. Example 3.18 computes the ideal defining the Rees algebra $\mathcal{R}(I)$ of $I = \langle t^4, s^2t^2, s^4 - st^3 \rangle$. The Macaulay2 package ReesAlgebra by David Eisenbud [**140**] allows us to compute $\mathcal{R}(I)$ as follows (some output suppressed):

```
i1 : R=ZZ/31991[s,t];

i2 : I=ideal(t^4,s^2*t^2,s^4-s*t^3);

i3 : syz(gens I)
o3 = {4} | s2  -st |
     {4} | -t2 s2  |
     {4} | 0   -t2 |

              3       2
o3 : Matrix R  <--- R

i4 : RI=reesIdeal I;

o4 : Ideal of R[w , w , w ]
                 0   1   2
```

```
i5 : transpose gens RI

o5 = {-1, -6}  | w_0st-w_1s2+w_2t2                       |
     {-1, -6}  | w_0s2-w_1t2                             |
     {-2, -9}  | w_0w_1t-w_1^2s+w_0w_2s                  |
     {-2, -9}  | w_0^2s-w_1^2t+w_0w_2t                   |
     {-4, -16} | w_0^3w_1-w_1^4+2w_0w_1^2w_2-w_0^2w_2^2 |

i6 : isLinearType RI

o6 = false
```

The two syzygies on I appear as the first two generators for the ideal defining $\mathcal{R}(I)$, and we see that $\mathrm{Sym}_R(I)$ and $\mathcal{R}(I)$ differ. ◁▷

It is a useful exercise (or see [**139**, Exercise A2.6]) to show that if R is a domain, then the kernel of the surjection from $\mathrm{Sym}_R(I)$ to $\mathcal{R}(I)$ consists of those elements of $\mathrm{Sym}_R(I)$ which are annihilated by R: the kernel of $\mathrm{Sym}_R(I) \to \mathcal{R}(I)$ is the R-torsion of the symmetric algebra.

Local Cohomology and the Jacobian Dual. In general the linear type condition is rare: $\mathrm{Sym}_R(I)$ typically has R-torsion. Busé and Jouanolou [**64**] proved that if the base locus is a zero-dimensional local complete intersection, then the R-torsion is rather mild, which can be phrased in terms of *local cohomology*.

Definition 3.37. For an ideal I and R-module M, the zeroth local cohomology is

$$H^0_I(M) = \{m \in M \mid I^j \cdot m = 0 \text{ for some } j \in \mathbb{N}\}.$$

Letting the role of I be played by the ideal $\mathfrak{m} = \langle x_1, \ldots, x_n \rangle$, there is a standard four term exact sequence in local cohomology (Theorem A4.1 of [**139**]):

$$0 \longrightarrow H^0_{\mathfrak{m}}(M) \longrightarrow M \longrightarrow \bigoplus_{i \in \mathbb{Z}} H^0(\widetilde{M}(i)) \longrightarrow H^1_{\mathfrak{m}}(M) \longrightarrow 0.$$

If we use $\mathrm{Sym}_R(I)$ for M in the above sequence, then if I is linear type outside $\mathfrak{m}$, this means $\mathcal{R}(I)$ and $\mathrm{Sym}_R(I)$ define the same sheaf outside $\mathbf{V}(\mathfrak{m})$, and $\mathcal{R}(I)$ is the third term in the sequence above. The surjection from $\mathrm{Sym}_R(I)$ to $\mathcal{R}(I)$ forces $H^1_{\mathfrak{m}}(\mathrm{Sym}_R(I))$ to vanish, and so we obtain the exact sequence

$$0 \longrightarrow H^0_{\mathfrak{m}}(\mathrm{Sym}_R(I)) \longrightarrow \mathrm{Sym}_R(I) \longrightarrow \mathcal{R}(I) \longrightarrow 0. \tag{3.37}$$

Recall that the *annihilator* $\mathrm{Ann}(M)$ of an R-module M is the ideal

$$\{r \in R \mid r \cdot m = 0 \text{ for all } m \in M\}.$$

A pair of theorems due to Busé and Jouanolou [**64**] are the key to using syzygies to study implicitization. Regarding $\mathcal{R}(I)$ as a graded R-module, here is the first.

Theorem 3.38. *For $\phi : S \to R$ from* (3.36), $\ker(\phi) = \mathrm{Ann}_S(\mathcal{R}(I)_0)$.

Proof. First, note that

$$\ker(\phi) = K \cap S,$$

where $K = \ker(\beta : R \otimes S \to \mathcal{R}(I) \subseteq R[e])$. This follows since if $F \in S$ satisfies

$$F(f_1 e, \ldots, f_d e) = 0 \in R[e],$$

then specializing to $e = 1$ shows $K \cap S \subseteq \ker(\phi)$. On the other hand, since the f_i are homogeneous, this means that $\ker(\phi)$ is also homogeneous, so $F(f_1 e, \ldots, f_d e) = e^m \cdot F(f_1, \ldots, f_d) = 0$, which implies $\ker(\phi) \subseteq K \cap S$. Since $\mathcal{R}(I) \simeq R \otimes S/K$, $F \in S$ annihilates the degree zero (in R) component of $\mathcal{R}(I)$ exactly if

$$F \cdot 1 \in K,$$

which holds exactly when $F \in K \cap S = \ker(\phi)$. □

Caveat: if we do not assume that R is a polynomial ring, then an additional hypothesis, that $H^0_{\mathfrak{m}}(R) = 0$, is necessary for the theorem to hold. Similar arguments yield the following.

THEOREM 3.39. *If $H^0_{\mathfrak{m}}(\mathrm{Sym}_R(I))_b = 0$ for all $b \geq a$, then*

$$\mathrm{Ann}_S(\mathrm{Sym}_R(I)_b) \subseteq \ker(\phi) \quad \text{for all } b \geq a,$$

and equality holds if I is of linear type outside $\mathfrak{m}$.

If I is not of linear type, then the next best case is when the defining ideal of $\mathcal{R}(I)$ has the *expected form*. To explain this terminology, we need the notion of the *Jacobian dual*.

DEFINITION 3.40. Let N denote the matrix of first syzygies on the ideal $I = \langle f_1, \ldots, f_d \rangle \subseteq R = k[x_1, \ldots, x_m]$, and write

$$[y_1, \ldots, y_d] \cdot N = [x_1, \ldots, x_m] \cdot B,$$

where the entries of B are linear in the y_i variables. The matrix B is the Jacobian dual of N. If the defining ideal of $\mathcal{R}(I)$ is generated by $[y_1, \ldots, y_d] \cdot N$ and the $d \times d$ minors of B, then $\mathcal{R}(I)$ is said to be of *expected form*.

The mystery matrix B that appeared in Example 3.35 is exactly the Jacobian dual. In Theorem 3.39, the hypothesis that I is linear type outside $\mathfrak{m}$ means that in high enough degree b in the R variables,

$$H^0_{\mathfrak{m}}(\mathrm{Sym}_R(I))_b = 0$$

and hence in high enough degree, $\mathrm{Sym}_R(I)$ and $\mathcal{R}(I)$ agree. Theorem 3.39 provides a path to connect implicitization problems to Fitting ideals and the MacRae invariant, and as we shall see, gives a concrete means to efficiently compute the implicit equation F. Example 3.35 is of expected form; Example 3.52 below is not.

Fitting Ideals and the Annihilator of a Module. A standard construction in algebra is that of the Fitting ideal of a module. For us, the context is that of a polynomial ring R and finitely generated R-module M. Then

DEFINITION 3.41. Given a presentation

$$R^p \xrightarrow{\psi} R^q \longrightarrow M \longrightarrow 0.$$

for M, $\mathrm{Fitt}_i(M)$ is the ideal of $q - i$ by $q - i$ minors of ψ; in particular $\mathrm{Fitt}_0(M)$ is the ideal of $q \times q$ minors of ψ.

It is a good exercise to show that the Fitting ideals of M are independent of choice of ψ; they capture information about the annihilator of the module.

THEOREM 3.42. $\mathrm{Fitt}_0(M) \subseteq \mathrm{Ann}(M)$.

PROOF. If $p < q$, then $\mathrm{Fitt}_0(M) = 0$ since ψ has no $q \times q$ minors. So assume $p \geq q$. Choose a $q \times q$ submatrix A of ψ such that $\det(A) \neq 0$, yielding a commuting diagram

$$\begin{array}{ccccccc} R^q & \xrightarrow{A} & R^q & \longrightarrow & M' & \longrightarrow & 0 \\ \downarrow & & \downarrow = & & \downarrow & & \\ R^p & \xrightarrow{\psi} & R^q & \longrightarrow & M & \longrightarrow & 0. \end{array}$$

Since M is a quotient of M', relations on M' are also relations on M and

$$\mathrm{Ann}(M') \subseteq \mathrm{Ann}(M).$$

We need to show that $\det(A)$ kills M'. If $a \in M' = \mathrm{coker}(A)$ and $b \in R^q$ maps to a, then

$$\det(A) \cdot a = 0 \iff \det(A) \cdot b = Ax \text{ for some } x \in R^q,$$

and such an x can be found by Cramer's rule. □

By Proposition 20.7 of [**139**], if M can be generated by r elements, then we have that $\mathrm{Ann}(M)^r \subseteq \mathrm{Fitt}_0(M)$. This implies that the radical ideals satisfy

$$\sqrt{\mathrm{Fitt}_0(M)} = \sqrt{\mathrm{Ann}(M)}.$$

Theorem 3.38 and Theorem 3.39 show that finding the implicit equation is a problem of computing annihilators, hence Fitting ideals will play a role.

Free Resolutions and the MacRae Invariant.

DEFINITION 3.43. A sequence of modules C_i and morphisms d_i

$$\mathcal{C}: \quad \cdots \longrightarrow C_{i+1} \xrightarrow{d_{i+1}} C_i \xrightarrow{d_i} C_{i-1} \xrightarrow{d_{i-1}} \cdots$$

is a *complex* if $\mathrm{im}(d_{i+1}) \subseteq \ker(d_i)$ for every i. It is *exact* at position i if $\mathrm{im}(d_{i+1}) = \ker(d_i)$. The i^{th} homology module is defined as $H_i(\mathcal{C}) = \ker(d_i)/\mathrm{im}(d_{i+1})$, hence $H_i(\mathcal{C}) = 0$ if and only if $\mathcal{C}$ is exact at position i.

An important special case occurs when the complex has the form

$$\mathcal{F}: \ 0 \longrightarrow F_n \xrightarrow{d_n} F_{n-1} \xrightarrow{d_{n-1}} \cdots \xrightarrow{d_2} F_1 \xrightarrow{d_1} F_0 \longrightarrow M \longrightarrow 0,$$

where the modules F_i are free of rank r_i, and the complex is everywhere exact. In this case $\mathcal{F}$ is called a *free resolution* of M; it is *minimal* if the d_i have no invertible entries.

By the Hilbert Syzygy Theorem (Corollary 19.7 of [**139**]), a finitely generated module M over the polynomial ring R has a finite free resolution, which stops in at most m steps, where m is the number of variables of R.

EXAMPLE 3.44. Consider the ideal $I = \langle x^2, y^2, z^2 \rangle \subseteq R = k[x, y, z]$. Then the free resolution of R/I takes the form

$$0 \longrightarrow R(-6) \xrightarrow{\begin{bmatrix} z^2 \\ -y^2 \\ x^2 \end{bmatrix}} R^3(-4) \xrightarrow{\begin{bmatrix} y^2 & z^2 & 0 \\ -x^2 & 0 & z^2 \\ 0 & -x^2 & -y^2 \end{bmatrix}} R^3(-2) \xrightarrow{\begin{bmatrix} x^2 & y^2 & z^2 \end{bmatrix}} R \longrightarrow R/I.$$

Notice that the first syzygies are spanned by the obvious relations: for every pair (f, g) in I, there is a relation $g \cdot f - f \cdot g = 0$. Such a syzygy is called a *Koszul*

syzygy, and the free resolution above has a simple construction in terms of exterior algebra.

DEFINITION 3.45. For any set of polynomials $\{f_1, \ldots, f_d\} \subseteq R$, the *Koszul complex* is given by

$$0 \longrightarrow \bigwedge^d R^d = R \longrightarrow \bigwedge^{d-1} R^d \longrightarrow \cdots \longrightarrow \bigwedge^2 R^d \xrightarrow{d_2} \bigwedge^1 R^d = R^d \xrightarrow{d_1} R,$$

with $d_i(e_{j_1} \wedge \cdots \wedge e_{j_i}) = \sum_{k=1}^{i} (-1)^k f_k e_{j_1} \wedge \cdots \wedge \widehat{e_{j_k}} \wedge \cdots \wedge e_{j_i})$. A straightforward computation shows that $d_{i-1} \circ d_i = 0$.

The Koszul complex is exact if and only if the f_i are a *regular sequence*: for all i, multiplication by f_i defines an injection on $R/\langle f_1, \ldots, f_{i-1}\rangle$. Geometrically, this means that each $\mathbf{V}(f_1, \ldots, f_{i-1})$ has codimension $i-1$ or equivalently intersecting $\mathbf{V}(f_1, \ldots, f_{i-1})$ with $\mathbf{V}(f_i)$ drops the dimension by exactly one. For this reason, a variety whose ideal is defined by a regular sequence is known as a *complete intersection*.

Now we turn to a special situation: let M be a module over a polynomial ring R, such that M is *torsion*: for each $m \in M$, there is some nonzero $r \in R$ such that $r \cdot m = 0$. Let $\mathcal{F}$ be a free resolution for M. Since R is a domain we may localize $\mathcal{F}$ at the zero ideal, yielding an exact sequence of vector spaces, with the d_i having entries in the field of fractions of R. Since the rank of F_n is r_n we may factor d_n as $\begin{bmatrix} \alpha_n \\ \beta_n \end{bmatrix}$, with $\operatorname{rank}(\alpha_n) = r_n$.

By exactness the image of d_n is the kernel of d_{n-1}, and so we may write d_{n-1} as

$$\begin{bmatrix} \beta_{n-1} & * \\ * & \alpha_{n-1} \end{bmatrix}$$

with β_{n-1} is of rank r_n and α_{n-1} is of rank $r_{n-1} - r_n$. Iterate this procedure. Since M is torsion if and only if it has rank zero, we also have that

$$\sum_{i=0}^{n} (-1)^i r_i = 0.$$

Thus in the factorization of $d_1 = [\beta_1 \mid \alpha_1]$, the rank of $\alpha_1 = r_1 - r_2 + r_3 - \cdots = r_0$.

DEFINITION 3.46. The *MacRae invariant* of M is

$$\mathcal{S}(M) = \frac{\det(\alpha_1) \cdot \det(\alpha_3) \cdots}{\det(\alpha_2) \cdot \det(\alpha_4) \cdots} = \prod \det(\alpha_i)^{(-1)^{i-1}}.$$

This definition seems to depend on the choice of the α_i. However, just as in the case of Fitting ideals, it turns out that the MacRae invariant is well-defined. For a proof of this, and the remarkable properties of the MacRae invariant below, we refer to the book of Northcott [**283**].

THEOREM 3.47. *The MacRae invariant is*

(1) *Independent of the choices of* α_i.
(2) *Multiplicative on short exact sequences.*
(3) *An element of R, rather than the field of fractions.*
(4) *Generator of the smallest principal ideal containing* $\operatorname{Fitt}_0(M)$.

If we hope to apply Theorem 3.47 above in conjunction with Theorem 3.38 or Theorem 3.39, we will need to obtain a resolution of $\mathcal{R}(I)_0$ or $\operatorname{Sym}_R(I)_b$ as an S-module. This takes us to our next key tool.

REMARK 3.48. If $\mathcal{F}_\bullet$ is not exact but simply a chain complex, it is possible to define the determinant of $\mathcal{F}_\bullet$, see [**165**]. We will not need this level of generality.

The Approximation Complex. The canonical example of a chain complex is the Koszul complex which appeared in Example 3.44. The Koszul complex takes center stage as we describe the work of Herzog, Simis and Vasconcelos [**202**], [**203**] on approximation complexes. Roughly speaking, the idea is to build a hybrid complex using $S = k[y_1, \ldots, y_d]$ and the ideal $I = \langle f_1, \ldots, f_d \rangle \subseteq R = k[x_1, \ldots, x_m]$.

DEFINITION 3.49. In the setting above, let d_i^f be the i^{th} Koszul differential on $\{f_1, \ldots, f_d\}$, Z_i the kernel of d_i^f, and d_i^S the Koszul differential on $\{y_1, \ldots, y_d\}$. The approximation complex $\mathcal{Z}$ is

$$\mathcal{Z}: \quad \cdots \longrightarrow Z_{i+1} \otimes_k S \xrightarrow{d_{i+1}^S} Z_i \otimes_k S \xrightarrow{d_i^S} Z_{i-1} \otimes_k S \xrightarrow{d_{i-1}^S} \cdots$$

Like the Koszul complex, $\mathcal{Z}$ depends only on I and not the choice of generators. To see the maps are well defined, note that $d^f d^S + d^S d^f = 0$. Let $\gamma \in \ker(d_i^f) = Z_i$, which implies $d^S d^f(\gamma) = 0$. Combining this with $(d^f d^S + d^S d^f)(\gamma) = 0$, we see that $d^f d^S(\gamma) = 0$; hence

$$d^S(\gamma) \in Z_{i-1} \otimes_k S.$$

Consider the rightmost homology of the approximation complex. Since Z_1 consists exactly of the syzygies on I, and Z_0 is R, we find

$$H_0(\mathcal{Z}) = R \otimes_k S/\langle \mathrm{Syz}(I) \rangle \simeq \mathrm{Sym}_R(I).$$

If $H_i(\mathcal{Z}) = 0$ for all $i \geq 1$, then the approximation complex will give a free resolution for $\mathrm{Sym}_R(I)$, and we will be able to apply the results on Fitting ideals and the MacRae invariant to obtain information about the implicit equation.

THEOREM 3.50. [**202**] *The approximation complex $\mathcal{Z}$ is acyclic if and only if*

$$f_{i+1} \cdot H_j(K(f_1, \ldots, f_i)) = 0 \text{ for all } i \in \{0, \ldots, d-1\},\ j > 0,$$

In [**57**], Busé and Chardin use a spectral sequence argument to prove that if $\mathbf{V}(I)$ is zero dimensional, then the approximation complex is acyclic outside $\mathbf{V}(\mathfrak{m})$ if and only if I is locally an *almost* complete intersection (ACI), that is, generated by at most one more equation than the codimension. In this case, Theorem 3.50 yields a free resolution of $\mathrm{Sym}_R(I)$. In the ACI case an extraneous factor appears in the MacRae invariant [**58**]; see Theorem 3.51 below.

Of course, we do not need the whole free resolution; by Theorem 3.39 we need the subresolution of $\mathrm{Sym}_R(I)_a$, where $H^0_{\mathfrak{m}}(\mathrm{Sym}_R(I)_b) = 0$ for all $b \geq a$. Busé and Jouanolou [**64**] and Busé and Chardin [**57**] show that when the base locus is empty, taking $a \geq (n-1)(m-1)$ suffices, where n is the degree of the f_i, and if the base locus $\mathcal{B}$ is nonempty, we can actually choose $a \geq (n-1)(m-1) - \varepsilon$, where ε is the minimal degree hypersurface containing the base locus.

Here is the precise result proved by Busé, Chardin and Jouanolou [**58**].

THEOREM 3.51. *If the base locus $\mathcal{B}$ is ACI, then for any $a \geq (n-1)(m-1) - \varepsilon$, the MacRae invariant of the approximation complex in degree a equals*

$$F^{\deg(\Phi)} \times \prod_{p \in \mathcal{B}} L_p^{e_p(I) - d_p(I)}$$

in the notation of (3.34)*. This simplifies to $F^{\deg(\Phi)}$ in the LCI case.*

These methods are used in the proof of Theorem 3.33 in Section 3.2.

In Example 3.35, we had $n = 3 = m$, so if I were basepoint free we would need the degree $(3-1)(3-1) = 4$ piece of the approximation complex. However, since the minimal degree curve through the six base points is a cubic, it suffices to look at the degree $4 - 3 = 1$ graded piece of $\mathrm{Sym}_R(I)$. This explains what happened in Example 3.35; the key takeaway is that *basepoints make the computation simpler!*

Our next example also has a local complete intersection base locus, but illustrates that typically we need more than two steps of the approximation complex to determine the implicit equation.

EXAMPLE 3.52. The ideal $I = \langle x_1^2, x_2^3 \rangle$ is a local complete intersection, and since I has a generator in degree two we need the degree two piece of the approximation complex. The map ϕ is determined by the cubics in I: $\begin{bmatrix} x_1^3 & x_1^2x_2 & x_2^3 & x_1^2x_3 \end{bmatrix}$ whose syzygies are generated by the columns of

$$\begin{bmatrix} -x_2 & -x_3 & 0 & 0 \\ x_1 & 0 & -x_3 & -x_2^2 \\ 0 & 0 & 0 & x_1^2 \\ 0 & x_1 & x_2 & 0 \end{bmatrix},$$

which has degree two component generated by the image of

$$\begin{bmatrix} -x_1x_2 & -x_2^2 & -x_2x_3 & -x_1x_3 & -x_2x_3 & -x_3^2 & 0 & 0 & 0 \\ x_1^2 & x_1x_2 & x_1x_3 & 0 & 0 & 0 & -x_2x_3 & -x_3^2 & -x_2^2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x_1^2 \\ 0 & 0 & 0 & x_1^2 & x_1x_2 & x_1x_3 & x_2^2 & x_2x_3 & 0 \end{bmatrix}.$$

Multiplying $[y_1, y_2, y_3, y_4]$ against M yields nine elements

$$[-y_1x_1x_2 + y_2x_1^2, \ldots, -y_2x_2^2 + y_3x_1^2],$$

and contracting against the quadrics in $\{x_1, x_2, x_3\}$ shows that $(Z_1)_2$ is generated by the columns of

$$M((Z_1)_2) = \begin{bmatrix} y_2 & 0 & 0 & y_4 & 0 & 0 & 0 & 0 & y_3 \\ -y_1 & y_2 & 0 & 0 & y_4 & 0 & 0 & 0 & 0 \\ 0 & 0 & y_2 & -y_1 & 0 & y_4 & 0 & 0 & 0 \\ 0 & -y_1 & 0 & 0 & 0 & 0 & y_4 & 0 & -y_2 \\ 0 & 0 & -y_1 & 0 & -y_1 & 0 & -y_2 & y_4 & 0 \\ 0 & 0 & 0 & 0 & 0 & -y_1 & 0 & -y_2 & 0 \end{bmatrix}.$$

Next, we need the term Z_2 in the approximation complex. There are two ways to do this: on the one hand, because I is a local complete intersection, we could simply compute the kernel of the matrix above. Or we could consider the kernel of the second Koszul differential on I, which is generated by the columns of

$$\begin{bmatrix} x_3 & x_2^3 & 0 & 0 \\ 0 & -x_1^2x_2 & 0 & x_1^2x_3 \\ 0 & x_1^3 & x_1^2x_3 & 0 \\ -x_2 & 0 & 0 & -x_2^3 \\ x_1 & 0 & -x_2^3 & 0 \\ 0 & 0 & x_1^2x_2 & x_1^3 \end{bmatrix}.$$

The quadratic component of the image is generated by

$$x_3(e_1 \wedge e_2) - x_2(e_1 \wedge e_4) + x_1(e_2 \wedge e_4).$$

Pushing this forward via the Koszul differential yields

$$x_3(x_1^3e_2 - x_1^2x_2e_1) - x_2(x_1^3e_4 - x_1^2x_3e_1) + x_1(x_1^2x_2e_4 - x_1^2x_2e_2).$$

Multiplying this by $\{x_1, x_2, x_3\}$ to obtain the quadratic component and expressing the result in terms of the ordered basis above shows that Z_2 has quadratic component spanned by the columns of

$$\begin{bmatrix} y_4 & 0 & 0 \\ 0 & y_4 & 0 \\ -y_1 & 0 & y_4 \\ -y_2 & 0 & 0 \\ y_1 & -y_2 & 0 \\ 0 & 0 & -y_2 \\ 0 & y_1 & 0 \\ 0 & 0 & y_1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Now that we have the approximation complex, we compute the MacRae invariant to get the implicit equation via Theorem 3.51. The determinant of the topmost 3×3 block of this is y_4^3, and the determinant of the complementary (hence, rightmost) block of the matrix $M((Z_1)_2)$ is $y_4^3 \cdot (y_2^3 - y_1^2y_3)$, so the implicit equation is given by

$$\frac{y_4^3 \cdot (y_2^3 - y_1^2y_3)}{y_4^3} = y_2^3 - y_1^2y_3,$$

which is easily checked to be correct. ◁▷

Multigraded Implicitization. A case which has attracted much interest in geometric modeling occurs when the ring R is multigraded. The hypersurface case which is the focus here was studied by Botbol [**39**], and we now describe this in more detail.

The most familiar example of a multigraded map occurs when the map ϕ is given by monomials. In this case the image is a *toric variety* and it is possible to compute the implicit equation (or equations if the image is not a hypersurface) using simplicial homology. Let the monomials have exponent vectors

$$\mathcal{A} = \{\mathbf{a}_1, \ldots, \mathbf{a}_d\} \subseteq \mathbb{Z}^n$$

and let A be the matrix with i^{th} column $\mathbf{a}_i$. If $Y_{\mathcal{A}} \subseteq k^d$ is the affine toric variety parametrized by the monomials, then [**106**, Proposition 1.1.9] tells us that vanishing ideal of $Y_{\mathcal{A}}$ in $k[y_1, \ldots, y_d]$ is the *toric ideal*

$$I_{\mathcal{A}} = \langle y^\alpha - y^\beta \mid \alpha, \beta \in \mathbb{N}^d \text{ and } \alpha - \beta \in \ker(A)\rangle \subseteq S = k[y_1, \ldots, y_d].$$

The ring $k[y_1, \ldots, y_d]$ is multigraded by $\mathbb{Z}^n$, where the degree of a monomial y^α is $A\alpha \in \mathbb{Z}^n$. The definition of $I_{\mathcal{A}}$ guarantees that $I_{\mathcal{A}}$ is homogeneous with respect to this multigrading. The graded piece in degree $\mathbf{m} \in \mathbb{Z}^n$ is denoted $(I_{\mathcal{A}})_{\mathbf{m}}$.

To any $\mathbf{m} \in \mathbb{Z}^n$ we associate a simplicial complex

$$\Delta_{\mathbf{m}} = \{J \subseteq \{1, \ldots, d\} \mid \mathbf{m} - \sum_{i \in J} \mathbf{a}_i \text{ lies in } \mathbb{N}\mathcal{A}\}.$$

A result of Hochster [**208**] shows that

$$\widetilde{H}_0(\Delta_{\mathbf{m}}, k) = \dim (I_{\mathcal{A}})_{\mathbf{m}}.$$

In particular, the generators of the ideal $I_{\mathcal{A}}$ are given by the zeroth reduced homology of certain simplicial complexes built from the semigroup $\mathbb{N}\mathcal{A}$.

EXAMPLE 3.53. Let $\mathcal{A} \subseteq \mathbb{Z}^2$ be generated by

$$\begin{bmatrix} 3 & 2 & 1 & 0 \\ 0 & 1 & 2 & 3 \end{bmatrix}.$$

We think of these lattice points as vertices $v_1, \ldots, v_4$. Then a computation shows that for

$$\mathbf{m} \in \left\{ \binom{2}{4}, \binom{3}{3}, \binom{4}{2} \right\},$$

we have $\widetilde{H}_0(\Delta_{\mathbf{m}}, k) \simeq k$. For example $\Delta_{\binom{4}{2}}$ has vertices v_1, v_2, v_3 and edge v_1v_3. Computing, we see that the ideal $I_{\mathcal{A}}$ has three generators

$$\begin{array}{lll} y_2^2 - y_1y_3 & \text{in degree} & \binom{4}{2} \\ y_1y_4 - y_2y_3 & \text{in degree} & \binom{3}{3} \\ y_3^2 - y_2y_4 & \text{in degree} & \binom{2}{4}. \end{array}$$

Hochster's result is more general, and shows that the higher syzygies of $I_{\mathcal{A}}$ may also be computed in similar fashion: first syzygies correspond to elements of $\widetilde{H}_1(\Delta_{\mathbf{m}}, k)$, and so on. In our example, a computation shows that

$$\widetilde{H}_1(\Delta_{\mathbf{m}}, k) = k$$

if and only if

$$\mathbf{m} \in \left\{ \binom{5}{4}, \binom{4}{5} \right\}, \text{ so there are exactly two first syzygies on } I_{\mathcal{A}}.$$

For more details and an in-depth exposition, see [**274**]. ◁▷

In the more general setting of a multigraded ideal I which is not generated by monomials, the approximation complex machinery works by taking a sufficiently high multidegree subcomplex of $\mathcal{Z}$. The multigraded structure means that the local cohomology must be computed with respect to a different ideal.

Recall that in geometric modeling, the most common surfaces are *triangular* surfaces, which correspond geometrically to $\mathbb{P}^2$, and *tensor product surfaces*, which correspond geometrically to $\mathbb{P}^1 \times \mathbb{P}^1$. Algebraically, a tensor product surface comes from a bigrading on $R = k[s, t, u, v]$, with s, t of degree $(1, 0)$ and u, v of degree $(0, 1)$. To make multigraded implicitization concrete, we focus on tensor product surfaces.

Let $R_{m,n}$ denote the graded piece of R in bidegree (m, n). A regular map $\mathbb{P}^1 \times \mathbb{P}^1 \xrightarrow{\Phi} \mathbb{P}^3$ is defined by four polynomials

$$\{p_0, p_1, p_2, p_3\} \subseteq R_{m,n}$$

with no common zeros on $\mathbb{P}^1 \times \mathbb{P}^1$. The empty locus on $\mathbb{P}^1 \times \mathbb{P}^1$ is defined by the bigraded ideal

$$\mathbf{m} = \langle s, t \rangle \cap \langle u, v \rangle = \langle su, sv, tu, tv \rangle.$$

Therefore the map Φ is basepoint free map if and only if the ideal

$$I = \langle p_0, p_1, p_2, p_3 \rangle$$

satisfies $\sqrt{I} = \mathbf{m}$.

EXAMPLE 3.54. Consider the ideal

$$I = \langle s^2u, s^2v, t^2u, t^2v + stv \rangle,$$

generated by four elements of bidegree $(2,1)$. The syzygies on I are generated by the columns of

$$\begin{bmatrix} -v & -t^2 & 0 & 0 & -tv \\ u & 0 & -st - t^2 & 0 & 0 \\ 0 & s^2 & 0 & -sv - tv & -sv \\ 0 & 0 & s^2 & tu & su \end{bmatrix},$$

which we encode as

$$\begin{array}{rcl} uy_2 - vy_1 & = & 0 \\ s^2y_3 - t^2y_1 & = & 0 \\ s^2y_4 - (st + t^2)y_2 & = & 0 \\ tuy_4 - (sv + tv)y_3 & = & 0 \\ suy_4 - svy_3 - tvy_1 & = & 0. \end{array}$$

If we were in the singly graded case, we would need to consider degree 2, and a basis for $\mathcal{Z}_1^2$ consists of $\{s,t,u,v\} \cdot uy_2 - vy_1$, and the remaining four relations. With respect to the ordered basis $\{s^2, st, t^2, su, sv, tu, tv, u^2, uv, v^2\}$ for R_2 the matrix for $d_1^2 : \mathcal{Z}_1^2 \longrightarrow \mathcal{Z}_0^2$ is

$$\begin{bmatrix} 0 & 0 & 0 & 0 & y_3 & y_4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -y_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & -y_1 & -y_2 & 0 & 0 \\ y_2 & 0 & 0 & 0 & 0 & 0 & 0 & y_4 \\ -y_1 & 0 & 0 & 0 & 0 & 0 & -y_3 & -y_3 \\ 0 & y_2 & 0 & 0 & 0 & 0 & y_4 & 0 \\ 0 & -y_1 & 0 & 0 & 0 & 0 & -y_3 & -y_1 \\ 0 & 0 & y_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -y_1 & y_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -y_1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

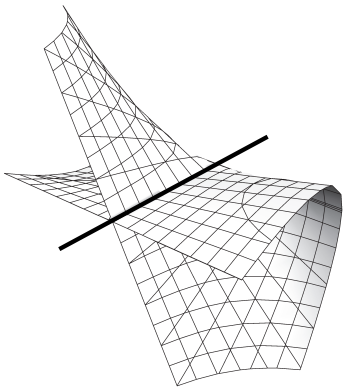
This matrix represents *all* the first syzygies of total degree two. The multigraded setting gives us extra structure to work with, and in Corollary 14 of [**42**], Botbol, Dickenstein and Dohm give a bound in terms of the multidegree; in the case of this example the bound tells us that it suffices to work with the subcomplex of the approximation complex of bidegree $(1,1)$. The first syzygies of bidegree $(1,1)$ correspond to the submatrix whose rows are indexed by $\{su, sv, tu, tv\}$, which is given by

$$\begin{bmatrix} y_2 & 0 & 0 & y_4 \\ -y_1 & 0 & -y_3 & -y_3 \\ 0 & y_2 & y_4 & 0 \\ 0 & -y_1 & -y_3 & -y_1 \end{bmatrix}.$$

The image of Φ is the determinantal hypersurface

$$H = \mathbf{V}(y_1y_2^2y_3 - y_2^2y_3^2 + 2y_1y_2y_3y_4 - y_1^2y_4^2)$$

pictured here:



In Theorems 3.19 and 3.20 in Section 3.2, we saw that the structure of the syzygy matrix is closely connected to the behavior of the singular locus of the implicit curve. This also holds for tensor product surfaces in $\mathbb{P}^3$ of bidegree $(2,1)$: in [**308**], Schenck, Seceleanu and Validashti show that there are six possible structures for the syzygy matrix, and describe the codimension one singular locus for each case. In our example above (from [**308**]) there is a unique syzygy of bidegree $(0,1)$, which in turn forces the codimension one singular locus of H to be the union of three lines $\mathbf{V}(y_1,y_3) \cup \mathbf{V}(y_2,y_4) \cup \mathbf{V}(y_1,y_2)$.

In what follows, the variables y_1, y_2, y_3, y_4 will be denoted w_3, w_1, w_2, w_0 for consistency with MACAULAY2 output. Since the generators of I are of bidegree $(2,1)$, the bidegree $(0,1)$ syzygy occurs in total degree $(2,2)$, and the variables w_i are of degree $(1,0,0)$. Therefore the generator $uy_2 - vy_1$ of the Rees ideal is now written $w_1u - w_3v$ and occurs in degree $(1,2,2)$, explaining the labelling of the last matrix below.

```
i1 : R=ZZ/31991[s,t,u,v,Degrees=>{{1,0},{1,0},{0,1},{0,1}}];

i2 : I=ideal(s^2*u,s^2*v,t^2*u,t^2*v+s*t*v);

i3 : rI=(res coker gens I).dd

          1                              4
o3 = 0 : R  <---------------------------- R  : 1
               | s2u t2u s2v stv+t2v |

          4                                          5
     1 : R  <------------------------------------------ R  : 2
               {2, 1} | -v -t2 0     0      0      |
               {2, 1} | 0  s2  0     -sv-tv -tv    |
               {2, 1} | u  0   st+t2 0      tu     |
               {2, 1} | 0  0   -s2   tu     -su+tu |

          5                          2
     2 : R  <---------------------- R  : 3
```

```
               {2, 2} | -t2 0   |
               {4, 1} | v   0   |
               {4, 1} | 0   -u  |
               {3, 2} | s-t -t  |
               {3, 2} | t   s+t |

          2
     3 : R  <----- 0 : 4
               0
i4 : transpose gens reesIdeal I

o4 = {-1, -2, -2} | w_1u-w_3v                                          |
     {-1, -4, -1} | w_2s2-w_3t2                                        |
     {-1, -3, -2} | w_0tu-w_2sv-w_2tv                                  |
     {-1, -3, -2} | w_0su-w_2sv-w_3tv                                  |
     {-1, -4, -1} | w_0s2-w_1st-w_1t2                                  |
     {-2, -5, -2} | w_1w_2t+w_0w_3s-w_0w_3t-w_1w_3t                    |
     {-2, -5, -2} | w_1w_2s-w_0w_3s+w_1w_3t                            |
     {-2, -4, -4} | w_0^2u2-2w_0w_2uv+w_2^2v2-w_2w_3v2                 |
     {-3, -6, -4} | w_1w_2^2v+w_0^2w_3u-2w_0w_2w_3v-w_1w_2w_3v         |
     {-4, -8, -4} | w_1^2w_2^2-2w_0w_1w_2w_3-w_1^2w_2w_3+w_0^2w_3^2|
```

Tensor product surfaces are a type of ruled surfaces, on which there is an extensive literature, ranging from the classical work of Edge [**138**] and Salmon [**305**] to contemporary work motivated by geometric modeling. We give pointers to relevant papers below. ◁▷

Directions for Future Research. There are a myriad of open questions at the interface of geometric modeling and algebraic geometry. A central problem is how to leverage results from situations where the free resolution of I is known to obtain results on the Rees algebra. For an example along these lines, if I is codimension two and Cohen-Macaulay, then I has a Hilbert-Burch resolution: $I = \langle f_1, \dots f_d \rangle$ is generated by the $(d-1) \times (d-1)$ minors of a $d \times d-1$ matrix, whose columns are the syzygies of I. In [**276**], Morey and Ulrich determine the equations for the Rees algebra, when all the syzygies are linear, and [**38**] gives an answer when exactly one syzygy is nonlinear; however the general case remains open. Some additional questions:

- Rees algebras of rational curves and connections to singularities. In addition to the works [**313**] and [**101**] mentioned earlier, there is a large literature on the interplay between syzygies and singularities of rational curves. In [**103**], Cox, Kustin, Polini and Ulrich carry out a comprehensive investigation; see also Cortadellas and D'Andrea [**90**] and [**91**].
- Tensor product surfaces have been extensively studied, but there remain many open questions. See [**122**], [**134**], [**142**], [**158**], [**371**] for some work on low degree cases; more generally [**39**], [**42**], [**119**] study toric implicitization.
- Matrix representations: rather than finding the actual implicit equation, focus on a matrix which drops rank at points of the implicitization. As Botbol,

Busé and Chardin write in [**40**] "matrices have to be seen as implicit representations on their own, without relying on the more classical implicit equation". For more work along these lines, see [**42**] and Theorem 3.34 in Section 3.2.

- Codimension three Gorenstein ideals have a structure theorem, and in [**248**], Kustin, Polini and Ulrich determine the equations for the Rees algebra. Is there a similar result for Gorenstein deviation two ideals, where there is a structure theory due to Huneke and Ulrich [**216**] and the resolutions are known from Kustin [**246**]?

- The resolution of ideals obtained from the submaximal minors of a matrix of variables may be constructed via representation theory and Bott-Borel-Weyl, see Weyman [**360**] for details. What can one say about the Rees algebra of such examples? About the simplest case of an Eagon-Northcott resolution?

- While Hochster's method provides a way to compute the implicit equations for the image of a monomial map, computing minimal generators for $\mathcal{R}(I)$ is nontrivial: see Lin [**259**] for a beautiful application of Alexander duality to the initial ideal of the maximal minors of a generic matrix, and Fouli and Lin [**156**] for additional results on squarefree monomial ideals.

- When $\mathcal{R}(I)$ is of quadratic type (determined by $\mathrm{Syz}(I)$ and $\mathrm{Syz}(I^2)$, [**59**] shows that half the degrees where the torsion is nonzero are understood.

- Methods from tropical geometry: if one can determine the degree of the implicit equation of a hypersurface, then the problem of actually describing it is linear algebra; in [**130**] and [**339**] tropical geometry is used to shed light on the problem.

CHAPTER 4

Rigidity Theory

This fascinating topic involves the study of frameworks built from bars and joints in $\mathbb{R}^2$ and $\mathbb{R}^3$. Given such a framework, a key question is whether it is rigid, and if not, what motions are possible. Studying this and other questions involves a wonderful combination of Euclidean geometry, algebraic geometry, and combinatorics.

4.1. Geometry of Rigidity

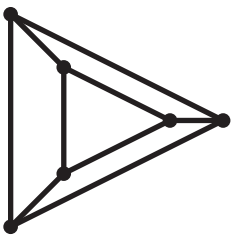
For most of this section, $G = (V, E)$ will be a simple graph with vertex set $V = [n] = \{1, \ldots, n\}$. An edge joining vertices $i \neq j$ is denoted $ij \in E$.

DEFINITION 4.1. A *bar-and-joint framework* in $\mathbb{R}^d$ for $G = (V, E)$ consists of:

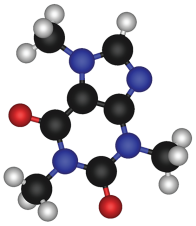
(1) Joints, which are points $\mathbf{q}_i \in \mathbb{R}^d$ for $i \in V$.
(2) Bars, which are line segments $\overline{\mathbf{q}_i\mathbf{q}_j}$ for $ij \in E$.

We often write a bar-and-joint framework as $\mathbf{q} = (\mathbf{q}_i)_{i \in V} \in (\mathbb{R}^d)^n$. The map $V \to \mathbb{R}^d$ sending $i \in V$ to $\mathbf{q}_i \in \mathbb{R}^d$ is typically injective. Frameworks where injectivity fails are said to be *degenerate*.

EXAMPLE 4.2. Here are two bar-and-joint frameworks:



Rigid but not infinitesimally rigid framework in $\mathbb{R}^2$



Caffeine molecule in $\mathbb{R}^3$

The framework on the left previews of some flavors of rigidity we will encounter. The molecule on the right can be modeled by a framework. Chemical properties of a molecule can change when folded (important in the study of proteins). This illustrates how rigidity (or lack thereof) arises naturally in applications. ◁▷

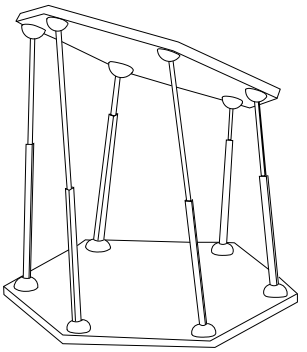
In a bar-and-joint framework, the bars meet at ball joints that can rotate in all directions. The idealized version presented in Definition 4.1 represents bars by line segments of zero thickness and joints by points also of zero thickness.

A related type of framework allows bars to meet at rigid bodies.

DEFINITION 4.3. A *body-and-bar framework* in $\mathbb{R}^d$ for $G = (V, E)$ consists of:
(1) Bodies, which are rigid bodies $B_i \subseteq \mathbb{R}^d$ for $i \in V$.
(2) Bars, which are line segments $\overline{\mathbf{q}_i\mathbf{q}_j}$ with $\mathbf{q}_i \in B_i$ and $\mathbf{q}_j \in B_j$ for an edge $e \in E$ connecting vertices $i, j \in V$.

In a body-and-bar framework, G is allowed to be a multigraph with no loops.

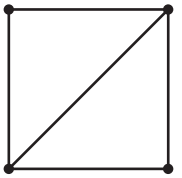
EXAMPLE 4.4. A well-known body-and-bar framework is the *Stewart-Gough platform*. Here is a picture from a paper by Bohigas, Maunbens and Ros [**35**]:



The graph G consists of two vertices connected by six edges. One rigid body (the base) is fixed to ground. By varying the lengths of the bars (the six legs), the other rigid body (the platform) can change orientation. Stewart-Gough platforms have many applications, including flight simulators and radio telescopes. ◁▷

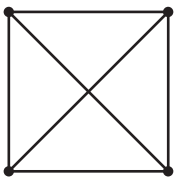
You will learn more about body-and-bar frameworks in Section 4.3. For now, we focus on the bar-and-joint case. Before giving a rigorous definition of rigidity, let's give some examples that rely on an intuitive notion of rigidity.

EXAMPLE 4.5. The rigidity of a framework can depend on the dimension of the ambient space. Consider the framework:

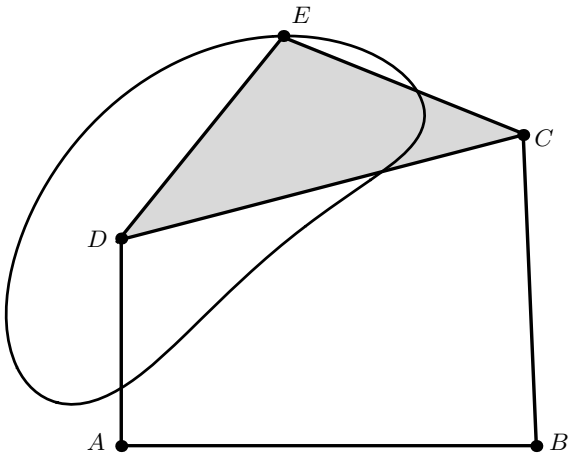


This is rigid in $\mathbb{R}^2$, but in $\mathbb{R}^3$, the two triangles can be swung back and forth along the diagonal. ◁▷

EXAMPLE 4.6. A bar-and-joint framework whose graph is the complete graph K_n is rigid in all dimensions. For example,



This framework has graph K_4 and can be regarded as lying in $\mathbb{R}^d$ for any $d \geq 2$. It is clearly rigid. Note that bars are allowed to intersect. ◁▷

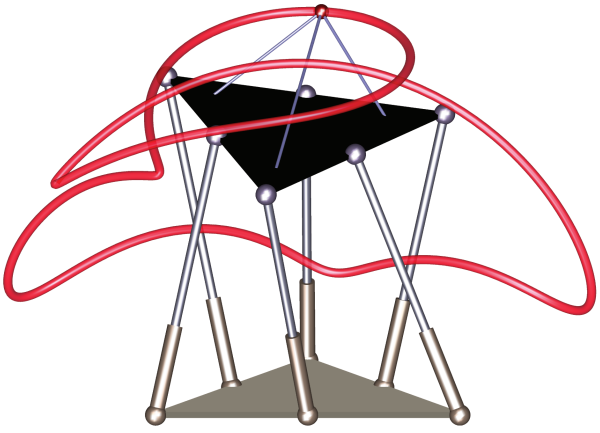


EXAMPLE 4.7. Here is an example based on the four-bar mechanism studied in Example 2.15 in Section 2.2: This framework consists the quadrilateral $ABCD$ with the triangle CDE attached to one edge. When we fix A and B and spin AD in a circle with center A, the point E traces out the curve show above. We found its equation in (2.21). ◁▷

EXAMPLE 4.8. A generic Stewart-Gough platform from Example 4.4 with *fixed edge lengths* is rigid. In this case, one can show that there are at most 40 platforms having the same lengths. See, for example, [**21**, Example 6.3].

However, motion is possible in special cases. A *Griffis-Duffy Type I platform* is a Stewart-Gough platform where the base and upper body are equilateral triangles with legs attached at the vertices and midpoints of the edges as indicated here:

(4.1)



With fixed edge lengths, there is now one degree of freedom, which gives the curve shown in red. This picture was created by Doug Arnold and Charles Wampler for the 2006-07 program on *Applications of Algebraic Geometry* at the Institute for Mathematics and its Applications. The image (4.1) also appears on the cover of the SIAM Journal on Applied Algebra and Geometry shown on page 95. ◁▷

Motion of a Framework. Before we can define rigidity, we need to understand what it means for a framework to move. For simplicity, we focus on bar-and-joint frameworks $\mathbf{q} = (\mathbf{q}_i)_{i\in V}$ in $\mathbb{R}^d$ for the graph $G = (V, E)$.

Since we are only interested in motions that preserve the edge lengths $\ell_{ij} = \|\mathbf{q}_i - \mathbf{q}_j\|^2$ for $ij \in E$, we have the following definition.

DEFINITION 4.9. A *motion* of $\mathbf{q} = (\mathbf{q}_i)_{i\in V}$ in $\mathbb{R}^d$ is a differentiable family of frameworks $\mathbf{q}(t) = (\mathbf{q}_i(t))_{i\in V}$ for $0 \le t \le 1$ such that:

(1) $\ell_{ij} = \|\mathbf{q}_i(t) - \mathbf{q}_j(t)\|^2$ for all $ij \in E$ and all $t \in [0, 1]$.
(2) $\mathbf{q}_i(0) = \mathbf{q}_i$ for all $i \in V$.

However, some motions are boring. To define this precisely, we need to recall the isometry group of $\mathbb{R}^d$.

DEFINITION 4.10. The *Euclidean group* $E(d)$ consists of all isometries of $\mathbb{R}^d$. The *special Euclidean group* $SE(d)$ is the subgroup consisting of all orientation preserving isometries of $\mathbb{R}^d$.

It is well known that any element of $E(d)$ can be written $v \mapsto Av + b$, where $A \in O(n)$ is orthogonal and $b \in \mathbb{R}^d$ represents translation. This lies in $SE(d)$ if and only if $\det(A) = 1$, so that $SE(d) \subseteq E(d)$ has index 2. Furthermore,

$$\dim E(d) = \dim SE(d) = \dim SO(d) + \dim \mathbb{R}^d = \binom{d}{2} + d = \binom{d+1}{2}. \tag{4.2}$$

Subsets $S, T \subseteq \mathbb{R}^d$ are *congruent* if $S = \gamma(T)$ for some $\gamma \in E(d)$.

We now define a boring motion of a framework.

DEFINITION 4.11. A motion $(\mathbf{q}_i(t))_{i\in V}$ of a framework is *trivial* if there is a differentiable map $\gamma : [0, 1] \to SE(d)$ with $\gamma(0) = I_n$ and

$$\mathbf{q}_i(t) = \gamma(t)\big(\mathbf{q}_i\big)$$

for all $i \in V$.

Flexibility and Local Rigidity. Any framework in $\mathbb{R}^d$ clearly has many trivial motions. But are there more? This leads to the following definition.

DEFINITION 4.12. A framework in $\mathbb{R}^d$ is *flexible* if it has a nontrivial motion. Otherwise the framework is *locally rigid*.

Later in the section we will see why the adjective "locally" is relevant. We next give a lemma that is useful when thinking about local rigidity.

LEMMA 4.13. *Let $\mathbf{q} = (\mathbf{q}_i)_{i\in V}$ be a framework in $\mathbb{R}^d$ with graph $G = (V, E)$. If G has a subgraph $G' = (V', E')$ such that $(\mathbf{q}_i)_{i\in V'}$ is locally rigid with respect to the edges $ij \in E'$, then the following are equivalent:*

(1) *The whole framework is locally rigid.*
(2) *Any motion of $\mathbf{q}$ fixing $\mathbf{q}_i$ for all $i \in V'$ is trivial.*

Furthermore, if $(\mathbf{q}_i)_{i\in V'}$ affinely spans $\mathbb{R}^d$, then (1) and (2) are equivalent to:

(3) *Any motion of $\mathbf{q}$ fixing $\mathbf{q}_i$ for all $i \in V'$ fixes the whole framework.*

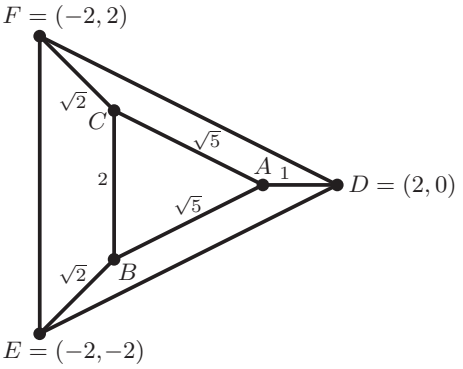
PROOF. (1) $\Rightarrow$ (2) is obvious. For (2) $\Rightarrow$ (1), suppose that we have a motion $(\mathbf{q}_i(t))_{i\in V}$ of the whole framework. Since $(\mathbf{q}_i)_{i\in V'}$ is locally rigid, there is a map $\gamma : [0, 1] \to SE(d)$ such that $\mathbf{q}_i(t) = \gamma(t)(\mathbf{q}_i)$ for $i \in V'$. Then $\bar{\mathbf{q}}_i(t)) = \gamma(t)^{-1} \circ \mathbf{q}_i(t)$

is a motion of the whole framework that fixes $\mathbf{q}_i$ for $i \in V'$. By (2), this is a trivial motion of the whole framework, and it follows easily that the original motion is also trivial.

Finally, assume $(\mathbf{q}_i(t))_{i\in V'}$ affinely spans $\mathbb{R}^d$. Elements of $SE(d)$ preserve affine spans, so the only way $\gamma(t) \in SE(d)$ can fix $(\mathbf{q}_i)_{i\in V'}$ is to be the identity. From here, (2) $\Rightarrow$ (3) follows easily, and then we are done since (3) $\Rightarrow$ (2) is obvious. □

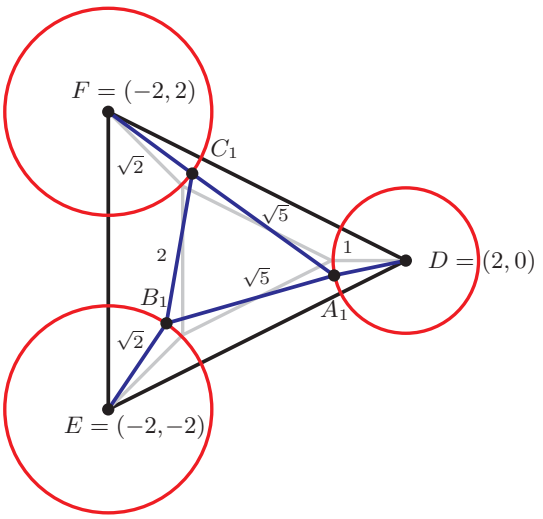
Example 4.14. We claim that the following framework is locally rigid:

(4.3)



This is one of the frameworks from Example 4.2.

The outer triangle is locally rigid since it comes from a complete graph. By Lemma 4.13, we need only consider motions that fix D, E, F. Since motions preserve lengths, the original inner triangle ABC (shown in gray below) must rotate along the red circles to a new inner triangle $A_1B_1C_1$ (shown in blue):



Write $A_1 = (x_1, y_1)$, $B_1 = (x_2, y_2)$, $C_1 = (x_3, y_3)$. Then preserving the six lengths A_1D, B_1E, C_1F and A_1B, B_1C, A_1C shows that x_i, y_i satisfy the six equations

$$(x_1-2)^2 + y_1^2 = 1,\ (x_2+2)^2 + (y_2+2)^2 = 2,\ (x_3+2)^2 + (y_3-2)^2 = 2$$
$$(x_1-x_2)^2 + (y_1-y_2)^2 = 5,\ (x_2-x_3)^2 + (y_2-y_3)^2 = 4,\ (x_1-x_3)^2 + (y_1-y_3)^2 = 5.$$

Eliminating y_1, x_2, y_2, x_3, y_3, we see that x_1 satisfies the equation

$$7743 - 5656x_1 + 890x_1^2 - 1272x_1^3 - 2105x_1^4 + 400x_1^5 = 0.$$

This quintic has roots $x_1 = 1, 5.8, -1.7697, 0.116098 \pm 1.36837i$. The only root that leads to real values of the other variables is $x_1 = 1$. Substituting $x_1 = 1$ into a lexicographic Gröbner basis and simplifying gives the polynomials

$$y_1^2, x_2 - y_1 + 1, y_1 + y_2 + 1, x_3 + y_1 + 1, y_1 + y_3 - 1. \tag{4.4}$$

These make it easy to see that $A_1 = A$, $B_1 = B$, $C_1 = C$. Thus any motion fixing D, E, F also fixes A, B, C, and rigidity follows by Lemma 4.13.

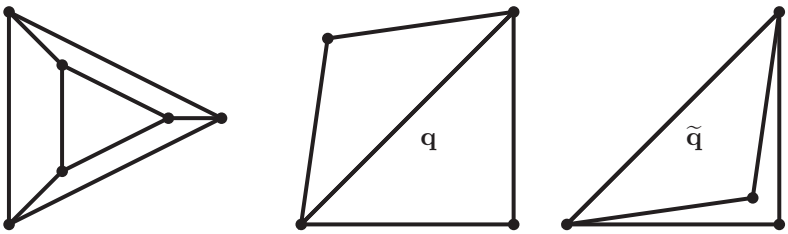
One interesting feature of (4.4) is the polynomial y_1^2, which from the point of view of scheme theory tells us that when $x_1 = 1$, we have an infinitesimal solution in the y_1 direction. This is related to the failure of (4.4) be infinitesimally rigid, as we will see below. ◁▷

Global Rigidity. Here is the strongest notion of rigidity.

DEFINITION 4.15. A framework $\mathbf{q}$ in $\mathbb{R}^d$ for $G = (V, E)$ is *globally rigid* if any framework $\widetilde{\mathbf{q}}$ in $\mathbb{R}^d$ for G with the same edge lengths as $\mathbf{q}$ is congruent to $\mathbf{q}$.

Any framework for the complete graph K_n is globally rigid in any dimension, and it is also easy to show that global rigidity implies local rigidity. The converse fails because even if a framework has only trivial motions, it might be able to jump to something different.

EXAMPLE 4.16. Consider the following frameworks in $\mathbb{R}^2$:



(a) (b)

Example 4.14 shows that framework (a) is locally rigid. The proof can be easily adapted to show global rigidity. On the other hand, the framework $\mathbf{q}$ in (b) is locally rigid in $\mathbb{R}^2$, but reflecting the upper left triangle about the diagonal (the "jump") gives the non-congruent framework $\widetilde{\mathbf{q}}$ with the same underlying graph and edge lengths. So $\mathbf{q}$ is not globally rigid. ◁▷

EXAMPLE 4.17. Examples 4.5 and 4.6 use an intuitive idea of rigidity. We now see that Example 4.6 is globally rigid in any dimension ≥ 2 since its graph is K_4.

On the other hand, Example 4.5 is locally rigid in $\mathbb{R}^2$ but not globally rigid since the lower right triangle is a degenerate framework with the same graph and edge lengths. To see why this makes sense, consider $\mathbf{q}$ and $\widetilde{\mathbf{q}}$ in Example 4.16(b). Deform the upper left vertex of $\mathbf{q}$ until the two triangles become congurent. As this happens, $\widetilde{\mathbf{q}}$ deforms to a degenerate framework supported on a triangle.

It follows that Example 4.5 and its lower right triangle are a limiting case of Example 4.16(b). This shows how degenerate frameworks can arise naturally. ◁▷

Infinitesimal Rigidity. By definition, motions of a framework are differentiable. Their derivatives will lead to a new flavor of rigidity.

Fix a framework $\mathbf{q} = (\mathbf{q}_i)_{i\in V}$ in $\mathbb{R}^d$ with $\ell_{ij} = \|\mathbf{q}_i - \mathbf{q}_j\|^2$ for $ij \in E$. Given a motion $\{\mathbf{q}_i(t)\}_{i\in V}$, its derivative at $t = 0$ is

$$\mathbf{q}_i' = \frac{d}{dt}\mathbf{q}_i(t)\Big|_{t\,=\,0}.$$

Motions preserve length, so that $\|\mathbf{q}_i(t)-\mathbf{q}_j(t)\|^2 = \ell_{ij}$ holds for all t. Differentiating this equation and setting $t = 0$, we get the system equations

$$(\mathbf{q}_i - \mathbf{q}_j)\cdot(\mathbf{q}_i' - \mathbf{q}_j') = 0, \quad ij \in E, \tag{4.5}$$

which are linear in the $\mathbf{q}_i'$.

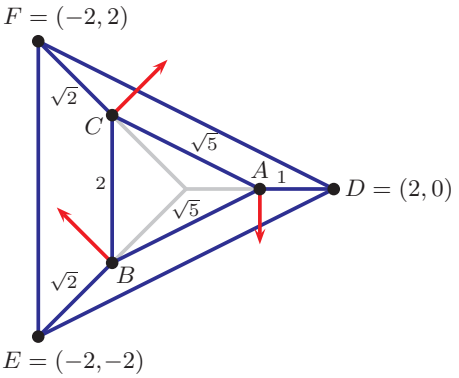
Definition 4.18. A framework $\mathbf{q}$ is *infinitesimally rigid* if the only solutions $(\mathbf{q}_i')_{i\in V}$ of (4.5) are those coming from trivial motions of $\mathbf{q}$.

In 1979, Asimow and Roth [**11**, §3] proved the following basic result.

Theorem 4.19. *Infinitesimal rigidity implies local rigidity.*

In general, the converse of Theorem 4.19 can fail, as shown by the following familiar example.

Example 4.20. Recall the framework analyzed in Example 4.14:



As before, we can fix D, E, F, so that $D' = E' = F' = 0$. Then (4.5) becomes:

$$\begin{aligned}
(A - D)\cdot A' = 0 &\implies A' = \alpha e_2\\
(B - E)\cdot B' = 0 &\implies B' = \beta(e_1 - e_2)\\
(C - F)\cdot C' = 0 &\implies C' = \gamma(e_1 + e_2)\\
(A - B)\cdot(A' - B') = 0 &\implies \alpha = \beta\\
(A - C)\cdot(A' - C') = 0 &\implies \alpha = -\gamma\\
(B - C)\cdot(B' - C') = 0 &\implies \beta = -\gamma,
\end{aligned}$$

where $\{e_1, e_2\}$ is the standard basis of $\mathbb{R}^2$. The solution $(\alpha, \beta, \gamma) = (-1, -1, 1)$ gives $A' = -e_2, B' = -e_1 + e_2, C' = e_1 + e_2$, shown above in red.

We conclude that this framework is locally rigid but not infinitesimally rigid. This happens because the lines $\overleftrightarrow{AD}, \overleftrightarrow{BE}, \overleftrightarrow{CF}$ intersect at a common point (shown above in gray). A generic framework on the same graph is infinitesimally rigid. ◁▷

The Rigidity Matrix. A nice aspect of infinitesimal rigidity is that it involves a system of linear equations (4.5). The matrix of this system will play an important role in what follows.

Suppose that the graph $G = (V, E)$ has $n = |V|$ vertices and $m = |E|$ edges. A framework $\mathbf{q} = (\mathbf{q}_i)_{i \in V}$ for G in $\mathbb{R}^d$ gives a linear map

$$T_{G,\mathbf{q}} : (\mathbb{R}^d)^n \longrightarrow \mathbb{R}^m$$

that maps velocity vectors $(\mathbf{q}'_1, \dots \mathbf{q}'_n) \in (\mathbb{R}^d)^n$ to the dot products

$$\big(\dots, (\mathbf{q}_i - \mathbf{q}_j) \cdot (\mathbf{q}'_i - \mathbf{q}'_j), \dots \big)_{ij \in E} \in \mathbb{R}^m.$$

The matrix $R_{G,\mathbf{q}}$ of $T_{G,\mathbf{q}}$ is the *rigidity matrix* of the framework $\mathbf{q}$. Note that the null space of $R_{G,\mathbf{q}}$ consists of the solutions of the system (4.5). This leads to our first result.

THEOREM 4.21. *If $\mathbf{q} = (\mathbf{q}_1, \dots, \mathbf{q}_n)$ affinely spans $\mathbb{R}^d$, then:*

(1) $\operatorname{rank}(R_{G,\mathbf{q}}) \leq dn - \binom{d+1}{2}$.
(2) $\mathbf{q}$ *is infinitesimally rigid if and only if* $\operatorname{rank}(R_{G,\mathbf{q}}) = dn - \binom{d+1}{2}$.

PROOF. The initial velocity vector of any motion of the framework gives an element of the kernel of $T_{G,\mathbf{q}}$. This includes trivial motions, so that

$$\operatorname{nullity}(R_{G,\mathbf{q}}) = \dim \ker(T_{G,\mathbf{q}}) \geq \dim SE(d) = \binom{d+1}{2},$$

where the middle inequality holds because $\mathbf{q}$ affinely spans $\mathbb{R}^d$. Then (1) follows from $\operatorname{nullity}(R_{G,\mathbf{q}}) + \operatorname{rank}(R_{G,\mathbf{q}}) = \dim (\mathbb{R}^d)^n = dn$. For (2), recall that infinitesimal rigidity means that all solutions of (4.5) come from trivial motions. Hence

$$\mathbf{q} \text{ is infinitesimally rigid} \iff \operatorname{nullity}(R_{G,\mathbf{q}}) = \dim SE(d) = \binom{d+1}{2}.$$

This translates to the rank condition in (2), and we are done. □

Since $R_{G,\mathbf{q}}$ has $m = |E|$ rows, we get the following nice corollary.

COROLLARY 4.22. *If $\mathbf{q} = (\mathbf{q}_1, \dots, \mathbf{q}_n)$ affinely spans $\mathbb{R}^d$, then:*

(1) $m \leq dn - \binom{d+1}{2}$ *if $R_{G,\mathbf{q}}$ has linearly independent rows.*
(2) $m \geq dn - \binom{d+1}{2}$ *if $\mathbf{q}$ is infinitesimally rigid.*

PROOF. In the situation of (1), $m = \operatorname{rank}(R_{G,\mathbf{q}})$, and then we are done by Theorem 4.21(1). In the situation of (2), we know that $\operatorname{rank}(R_{G,\mathbf{q}}) = dn - \binom{d+1}{2}$ by Theorem 4.21(2), which is bounded by the number of rows m of $R_{G,\mathbf{q}}$. □

The appearance of "linearly independent rows" in Corollary 4.22 hints at the matroidal aspect of rigidity. We will explore this further in the next section.

Other Flavors of Rigidity. So far, we seen three flavors of rigidity: local, global, and infinitesimal. There are also variants to consider:

- Each of these has a *minimal* version, where the property fails if any edge is removed.
- Each of these has a *generic* version. We will say more about this below.

We will also learn about *combinatorial rigidity* in Section 4.2.

Algebraic Geometry. It is now time to explore the algebraic geometry of rigidity. We begin with the *configuration variety*, which for $G = (V, E)$ parametrizes *all* frameworks in $\mathbb{R}^d$ with given edge lengths. As before, we set $n = |V|$ and $m = |E|$.

From G, we get the *length-squared map* $\ell_{G,d} : (\mathbb{R}^d)^n \to \mathbb{R}^m$ defined by

$$\ell_{G,d}(\mathbf{q}) = (\|\mathbf{q}_i - \mathbf{q}_j\|^2)_{ij \in E}.$$

This is a polynomial map since we are squaring the lengths. Then, given $\ell \in \mathbb{R}^m_{\geq 0}$, the *configuration variety* of G and ℓ is

$$C_{G,d,\ell} = \ell_{G,d}^{-1}(\ell) \subseteq (\mathbb{R}^d)^n. \tag{4.6}$$

This is a real algebraic variety. If $\mathbf{q}$ is a framework for G in $\mathbb{R}^d$, then $C_{G,d,\ell}$ for $\ell = \ell_{G,d}(\mathbf{q})$ consists of all frameworks for G with the same edge lengths as $\mathbf{q}$.

One easily shows that the Jacobian matrix of $\ell_{G,d}$ at $\mathbf{q} \in (\mathbb{R}^d)^n$ satisfies

$$\tfrac{1}{2} J\ell_{G,d}(\mathbf{q}) = R_{G,\mathbf{q}}, \tag{4.7}$$

so that up to a constant, the rigidity matrix $R_{G,\mathbf{q}}$ computes the derivative of the length-squared map. This has the following nice consequence for the associated configuration variety.

THEOREM 4.23. *Let* $\mathbf{q}$ *be a framework for* G *in* $\mathbb{R}^d$ *with* $\ell = \ell_{G,d}(\mathbf{q})$. *If* $R_{G,\mathbf{q}}$ *has rank* m, *then the configuration variety* $C_{G,d,\ell}$ *is smooth at* $\mathbf{q}$ *of dimension* $dn - m$.

PROOF. Thinking of $C_{G,d,\ell}$ as a real manifold, this follows from the Implicit Function Theorem since $J\ell_{G,d}(\mathbf{q})$ has rank m by (4.7). Translated to algebraic geometry, the rank condition implies smoothness of the desired dimension. □

Here is a fun application.

EXAMPLE 4.24. Let $P \subseteq \mathbb{R}^3$ be a 3-dimensional polytope. The edges of P give a framework $\mathbf{q}$ in $\mathbb{R}^3$ whose underlying graph will be denoted G. In 1978, Asimow and Roth [**10**] used results of Cauchy and Gluck to prove the following theorem.

THEOREM 4.25. *The edge framework* $\mathbf{q}$ *is locally rigid for* G *in* $\mathbb{R}^3$ *if and only if every face of* P *is a triangle.*

PROOF. First suppose that the faces of P are triangles. Since P is also convex, it follows that any small perturbation of the vertices gives a combinatorially equivalent polytope. Any motion $\mathbf{q}(t)$ starts from $\mathbf{q}$, so that for small values of t, $\mathbf{q}(t)$ will be the edge framework of a polytope P_t combinatorially equivalent to P. Since triangles are globally rigid, the corresponding faces of P and P_t will be congruent. Then an 1813 theorem of Cauchy [**68**] implies that P and P_t are congruent, and local rigidity follows.

For the other direction, assume that $\mathbf{q}$ is locally rigid in $\mathbb{R}^3$ and set $\ell = \ell_{G,3}(\mathbf{q})$. Ideas of Gluck [**172**] from 1975 imply that $J\ell_{G,3}(\mathbf{q})$ has rank m since $\mathbf{q}$ comes from a convex polytope (see [**298**, Theorem 6.3] for a careful proof). By Theorem 4.23, it follows that the configuration variety $C_{G,3,\ell}$ has dimension $3n - m$.

The affine span of $\mathbf{q}$ is $\mathbb{R}^3$ since $\dim(P) = 3$. Hence the $SE(3)$-orbit of $\mathbf{q}$ is smooth of dimension $\dim(SE(3)) = \binom{3+1}{2} = 6$. Local rigidity implies that this orbit coincides with $C_{G,3,\ell}$ near $\mathbf{q}$. By the previous paragraph, we conclude that

$$3n - m = 6, \text{ hence } m = 3n - 6.$$

To see why this forces the faces of P to be triangles, we use Euler's formula

$$f - m + n = 2,$$

which connects the number f of faces of P to the number m of edges and number n of vertices. Then we obtain

$$m = 3n - 6 = 3(n-2) = 3(m-f) = m + (2m - 3f),$$

which implies $2m - 3f = 0$. However, in the polytope P, every edge meets exactly two faces and every face has *at least* three edges. This tells us that $2m - 3f \geq 0$. The only way $2m - 3f \geq 0$ can become $2m - 3f = 0$ is for every face of P to have *exactly* three edges, i.e., a triangle. Thus P has triangular faces, as desired. □

The 1813 theorem of Cauchy mentioned above has an elegant proof that can be found in *Proofs from THE BOOK* [**6**]. ◁▷

The Cayley-Menger Matrix. When $G = K_n$, the image of the map

$$\ell_{K_n,d} : (\mathbb{R}^d)^n \to \mathbb{R}^{\binom{n}{2}} \tag{4.8}$$

can be described explicitly using equations and inequalities. We begin with an example of the equations involved.

Example 4.26. Three points $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3 \in \mathbb{R}^2$ determine a triangle whose area is given by Heron's formula

$$A = \sqrt{s(s-a)(s-b)(s-c)}$$

for edge lengths $a = \|\mathbf{q}_1 - \mathbf{q}_2\|, b = \|\mathbf{q}_1 - \mathbf{q}_3\|, c = \|\mathbf{q}_2 - \mathbf{q}_3\|$ and semiperimeter $s = \frac{1}{2}(a+b+c)$. After some algebra, one gets the equation

$$-16A^2 = \det \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & \|\mathbf{q}_1 - \mathbf{q}_2\|^2 & \|\mathbf{q}_1 - \mathbf{q}_3\|^2 \\ 1 & \|\mathbf{q}_2 - \mathbf{q}_1\|^2 & 0 & \|\mathbf{q}_2 - \mathbf{q}_3\|^2 \\ 1 & \|\mathbf{q}_3 - \mathbf{q}_1\|^2 & \|\mathbf{q}_3 - \mathbf{q}_2\|^2 & 0 \end{bmatrix}.$$

In particular, the determinant vanishes when the points $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$ lie on the horizontal axis in $\mathbb{R}^2$. It follows that since $\ell_{K_3,1} : (\mathbb{R}^1)^3 \to \mathbb{R}^3$ is defined by

$$\ell_{K_3,1}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) = (\|\mathbf{q}_1 - \mathbf{q}_2\|^2, \|\mathbf{q}_1 - \mathbf{q}_3\|^2, \|\mathbf{q}_2 - \mathbf{q}_3\|^2),\ \mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3 \in \mathbb{R}^1 = \mathbb{R},$$

the above determinant gives a polynomial that vanishes on the image of $\ell_{K_3,1}$. ◁▷

Cayley's first paper [**69**], published in 1841, generalized this result to higher dimensions. Suppose that we have n points $\mathbf{q} = (\mathbf{q}_1, \ldots, \mathbf{q}_n)$ in d-dimensional space $\mathbb{R}^d$. Then we have the $(n+1) \times (n+1)$ *Cayley-Menger matrix*

$$CM(\mathbf{q}) = \begin{bmatrix} 0 & 1 & 1 & 1 & \cdots & 1 \\ 1 & 0 & \|\mathbf{q}_1 - \mathbf{q}_2\|^2 & \|\mathbf{q}_1 - \mathbf{q}_3\|^2 & \cdots & \|\mathbf{q}_1 - \mathbf{q}_n\|^2 \\ 1 & \|\mathbf{q}_2 - \mathbf{q}_1\|^2 & 0 & \|\mathbf{q}_2 - \mathbf{q}_3\|^2 & \cdots & \|\mathbf{q}_2 - \mathbf{q}_n\|^2 \\ 1 & \|\mathbf{q}_3 - \mathbf{q}_1\|^2 & \|\mathbf{q}_3 - \mathbf{q}_2\|^2 & 0 & \cdots & \|\mathbf{q}_3 - \mathbf{q}_n\|^2 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \|\mathbf{q}_n - \mathbf{q}_1\|^2 & \|\mathbf{q}_n - \mathbf{q}_2\|^2 & \|\mathbf{q}_n - \mathbf{q}_3\|^2 & \cdots & 0 \end{bmatrix}.$$

The n points $\mathbf{q}$ determine an $(n-1)$-simplex in $\mathbb{R}^d$ whose $(n-1)$-dimensional

volume V_{n-1} satisfies

$$(-1)^n 2^{n-1}((n-1)!)^2 V_{n-1}^2 = \det CM(\mathbf{q})$$

(see [**34**, Section IV.40]). Example 4.26 is the special case when $n = 3$.

The volume V_{n-1} is positive if and only if $\mathbf{q}_1, \ldots, \mathbf{q}_n$ are affinely independent in $\mathbb{R}^d$. This is impossible if $d < n-1$, so that $\det CM(\mathbf{q}) = 0$, i.e., $\operatorname{rank} CM(\mathbf{q}) < n+1$ in this case. In fact, the dimension d puts a strong upper bound on $\operatorname{rank} CM(\mathbf{q})$.

THEOREM 4.27. *For $\mathbf{q} \in (\mathbb{R}^d)^n$, the matrix $CM(\mathbf{q})$ has rank $\leq d+2$, i.e., its $(d+3) \times (d+3)$ minors all vanish.*

PROOF. Given $\mathbf{q}_1, \ldots, \mathbf{q}_n \in \mathbb{R}^d$ and $2 \leq i, j \leq n$, the law of cosines for the triangle with edges $\mathbf{q}_i - \mathbf{q}_1$, $\mathbf{q}_j - \mathbf{q}_1$, $\mathbf{q}_i - \mathbf{q}_j$ implies that

$$(\mathbf{q}_i - \mathbf{q}_1) \cdot (\mathbf{q}_j - \mathbf{q}_1) = \tfrac{1}{2}\big(\|\mathbf{q}_i - \mathbf{q}_1\|^2 + \|\mathbf{q}_j - \mathbf{q}_1\|^2 - \|\mathbf{q}_i - \mathbf{q}_j\|^2\big). \tag{4.9}$$

These dot products form the *Gram matrix* of $\mathbf{q}_2 - \mathbf{q}_1, \ldots, \mathbf{q}_n - \mathbf{q}_1$, denoted $S_1(\mathbf{q})$ (the subscript 1 indicates subtraction by $\mathbf{q}_1$). The matrix $S_1(\mathbf{q})$ is symmetric of size $(n-1) \times (n-1)$.

To relate $CM(\mathbf{q})$ and $S_1(\mathbf{q})$, let

$$\begin{aligned} \ell_{ij} &= \|\mathbf{q}_i - \mathbf{q}_j\|^2, \qquad 1 \leq i, j \leq n \\ s_{ij} &= \tfrac{1}{2}\big(\ell_{i1} + \ell_{j1} - \ell_{ij}\big), \ 2 \leq i, j \leq n. \end{aligned}$$

Note that $\ell_{ij} = \ell_{ji}$ and $s_{ij} = s_{ji}$, and $s_{ij} = (\mathbf{q}_i - \mathbf{q}_1) \cdot (\mathbf{q}_j - \mathbf{q}_1)$ by (4.9). This allows us to write the Gram matrix as $S_1(\mathbf{q}) = (s_{ij})_{i,j=2}^n$. Then

$$-2S_1(\mathbf{q}) = (-2s_{ij})_{i,j=2}^n = (\ell_{ij} - \ell_{i1} - \ell_{1j})_{i,j=2}^n$$

leads to the following row and column operations on $CM(\mathbf{q})$:

$$\begin{bmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & 0 & \ell_{12} & \cdots & \ell_{1n} \\ 1 & \ell_{21} & 0 & \cdots & \ell_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \ell_{n1} & \ell_{n2} & \cdots & 0 \end{bmatrix} \overset{A}{\rightsquigarrow} \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & \ell_{12} & \cdots & \ell_{1n} \\ 0 & \ell_{21} & & & \\ \vdots & \vdots & & -2S_1(\mathbf{q}) & \\ 0 & \ell_{n1} & & & \end{bmatrix} \overset{B}{\rightsquigarrow} \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & & & \\ \vdots & \vdots & & -2S_1(\mathbf{q}) & \\ 0 & 0 & & & \end{bmatrix}.$$

We index the rows and columns of $CM(\mathbf{q})$ by $0, 1, \ldots, n$. The row and column operations $\overset{A}{\rightsquigarrow}$ subtract row and column 1 from rows and columns $2, \ldots, n$, and the row and column operations $\overset{B}{\rightsquigarrow}$ subtract suitable multiples of row and column 0 from rows and columns $2, \ldots, n$. Row and column operations preserve rank, so that

$$\operatorname{rank} CM(\mathbf{q}) = 2 + \operatorname{rank} S_1(\mathbf{q}). \tag{4.10}$$

Since $\mathbf{q}_2 - \mathbf{q}_1, \ldots, \mathbf{q}_n - \mathbf{q}_1 \in \mathbb{R}^d$ implies $\operatorname{rank} S_1(\mathbf{q}) \leq d$, the theorem follows. □

EXAMPLE 4.28. When $n = 3$ and $d = 1$, $CM(\mathbf{q})$ is the 4×4 matrix shown in Example 4.26. Then $\det CM(\mathbf{q}) = 0$ in this case follows from Theorem 4.27 since $d + 3 = 4$. ◁▷

The rank condition of Theorem 4.27 gives equations that vanish on the image of the map (4.8). However, there are also inequalities, since the Gram matrix $S_1(\mathbf{q})$ from (4.9) is *positive semidefinite* (PSD). Let's use this to characterize the image of $\ell_{K_n,d} : (\mathbb{R}^d)^n \to \mathbb{R}^{\binom{n}{2}}$. We will use $\ell_{ij} = \ell_{ji}$ for coordinates of $\mathbb{R}^{\binom{n}{2}}$, so $\ell_{K_n,d}$ is defined by $\ell_{ij} = \|\mathbf{q}_i - \mathbf{q}_j\|^2$ for $\mathbf{q} \in (\mathbb{R}^d)^n$.

THEOREM 4.29. *A point $\ell = (\ell_{ij}) \in \mathbb{R}^{\binom{n}{2}}$ is in the image of the length-squared map $\ell_{K_n,d} : (\mathbb{R}^d)^n \to \mathbb{R}^{\binom{n}{2}}$ if and only if*

(1) $$\operatorname{rank} \begin{bmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & 0 & \ell_{12} & \cdots & \ell_{1n} \\ 1 & \ell_{21} & 0 & \cdots & \ell_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \ell_{n1} & \ell_{n2} & \cdots & 0 \end{bmatrix} \leq d+2.$$

(2) *The $(n-1)\times(n-1)$ matrix $S = (s_{ij})_{i,j=2}^n$, $s_{ij} = \frac{1}{2}\big(\ell_{i1} + \ell_{j1} - \ell_{ij}\big)$, is* PSD.

PROOF. One direction follows from the above discussion. For the converse, let S be PSD of size $(n-1)\times(n-1)$. Then (1) and (4.10) imply that $\operatorname{rank} S \leq d$. Being PSD, S is the Gram matrix of $n-1$ vectors in $\mathbb{R}^{n-1}$, which we denote $\mathbf{q}_2, \ldots, \mathbf{q}_n$. Set $\mathbf{q}_1 = 0 \in \mathbb{R}^{n-1}$ and let $W = \operatorname{Span}(\mathbf{q}_2, \ldots, \mathbf{q}_n)$. The rank condition on S implies that $\dim W \leq d$, so we can regard $\mathbf{q}_1, \ldots, \mathbf{q}_n$ as vectors in $\mathbb{R}^d$. From here, it follows easily that $\ell_{ij} = \|\mathbf{q}_i - \mathbf{q}_j\|^2$ for $1 \leq i, j \leq n$. □

The name "Cayley-Menger" was chosen to honor Cayley's 1841 introduction of the matrix $CM(\mathbf{q})$ [**69**] and Menger's 1931 characterization of the image of $\ell_{K_n,d}$ using equations and inequalities [**270**]. The connection to PSD matrices via $s_{ij} = \frac{1}{2}\big(\ell_{i1} + \ell_{j1} - \ell_{ij}\big)$ is due to Schoenberg in 1935 [**309**].

The Cayley-Menger Variety. The equations and inequalities defining the image of $\ell_{K_n,d} : (\mathbb{R}^d)^n \to \mathbb{R}^{\binom{n}{2}}$ in Theorem 4.29 give equations that define the Zariski closure of the image in $\mathbb{R}^{\binom{n}{2}}$. This remains true over $\mathbb{C}$, which allows us to apply some classic results from algebraic geometry.

The defining equations of the image are easily seen to be homogeneous in the coordinates ℓ_{ij} of $\mathbb{C}^{\binom{n}{2}}$. Hence they define the *Cayley-Menger variety*

$$CM_{n,d,\mathbb{C}} \subseteq \mathbb{P}^{\binom{n}{2}-1}.$$

We compute the dimension and degree of $CM_{n,d,\mathbb{C}}$ as follows.

THEOREM 4.30. *The Cayley-Menger variety $CM_{n,d,\mathbb{C}} \subseteq \mathbb{P}^{\binom{n}{2}-1}$ has dimenion*

$$\dim CM_{n,d,\mathbb{C}} = dn - \tbinom{d+1}{2} - 1$$

and degree

$$\deg CM_{n,d,\mathbb{C}} = \prod_{\ell=0}^{n-d-2} \frac{\binom{n-1+\ell}{n-d-1-\ell}}{\binom{2\ell+1}{\ell}}.$$

PROOF. We follow Borcea and Streinu [**36**]. The change of variables $s_{ij} = \frac{1}{2}\big(\ell_{i1} + \ell_{j1} - \ell_{ij}\big)$ in Theorem 4.29 transforms the image of $\ell_{K_n,d}$ into the set of $(n-1)\times(n-1)$ symmetric PSD matrices of rank $\leq d$. Proposition 5.1.1 of [**201**] implies that the Zariski closure is the variety

$$\operatorname{Sym}_{n-1,d,\mathbb{R}} = \big\{(n-1)\times(n-1) \text{ real symmetric matrices of rank} \leq d\big\}.$$

This remains true over $\mathbb{C}$, so that projectively, the change of variables induces is isomorphism of projective varieties

$$CM_{n,d,\mathbb{C}} \simeq \mathbb{P}\operatorname{Sym}_{n-1,d,\mathbb{C}}, \tag{4.11}$$

where $\mathbb{P}\operatorname{Sym}_{n-1,d,\mathbb{C}} \subseteq \mathbb{P}^{\binom{n}{2}-1}$ is the projective version of $\operatorname{Sym}_{n-1,d,\mathbb{C}} \subseteq \mathbb{C}^{\binom{n}{2}}$.

The variety $\mathbb{P}\mathrm{Sym}_{n-1,d,\mathbb{C}}$ is a classic example of a determinantal variety. Its dimension and degree were computed by Gambelli in 1905 [**166**] and are given by the formulas in the statement of the theorem. Modern proofs are due Jósefiak, Lascoux and Pragacz in 1982 [**230**] and Harris and Tu in 1984 [**188**]. □

Back to Frameworks. The Cayley-Menger variety $CM_{n,d,\mathbb{C}}$ is lovely, but what does it have to do with rigidity? The map $\ell_{K_n,d}$ uses the complete graph K_n, whose frameworks are always rigid. How can this be interesting?

What saves the day is that K_n has lots of interesting subgraphs. In fact, any graph with n vertices is isomorphic to a subgraph of K_n. This simple observation has many consequences, including the *rigidity matroid* to be defined in Section 4.2 and a fun application of Cayley-Menger varieties due to Borcea and Streinu [**36**], which we will now explain.

We first set up some standard notation:

$$\begin{aligned} [n] &= \{1,\dots,n\} \\ \tbinom{[n]}{2} &= \big\{\{i,j\} \mid i,j \in [n],\, i \neq j\big\} = \big\{ij \mid i,j \in [n],\, i \neq j\big\}. \end{aligned} \tag{4.12}$$

Thus K_n has vertices $[n]$ and edges $\binom{[n]}{2}$, and an arbitrary graph G with n vertices can be represented as $G = ([n], E)$ for $E \subseteq \binom{[n]}{2}$. In this notation, the length-squared maps for K_n and G fit together to give the commutative diagram

$$\begin{array}{ccc} (\mathbb{R}^d)^n & \xrightarrow{\ell_{K_n,d}} & \mathbb{R}^{\binom{n}{2}} \\ & \searrow^{\ell_{G,d}} & \downarrow{\scriptstyle \pi_E} \\ & & \mathbb{R}^m \end{array} \tag{4.13}$$

where $m = |E|$ and π_E is projection onto the coordinates of $\mathbb{R}^{\binom{n}{2}}$ indexed by E. Given a framework $\mathbf{q}$ for $G = ([n], E)$ in $\mathbb{R}^d$, we define the affine subspace

$$L_{\mathbf{q}} = \pi_E^{-1}(\ell_{G,d}(\mathbf{q})) \subseteq \mathbb{R}^{\binom{n}{2}}.$$

Using this, we can describe congruence classes of frameworks for G using Cayley-Menger varieties as follows.

LEMMA 4.31. *Let $\mathbf{q}$ be a framework for $G = ([n], E)$ in $\mathbb{R}^d$. Then:*

(1) *If $\tilde{\mathbf{q}}$ is another framework for G in $\mathbb{R}^d$ with the same edge lengths as $\mathbf{q}$, then $\tilde{\mathbf{q}}$ is congruent to $\mathbf{q}$ if and only if $\ell_{K_n,d}(\tilde{\mathbf{q}}) = \ell_{K_n,d}(\mathbf{q})$*
(2) *There is a bijection*

$$\left\{\begin{array}{l}\text{congruence classes of frameworks for } G \\ \text{in } \mathbb{R}^d \text{ with the same edge lengths as } \mathbf{q}\end{array}\right\} \simeq L_{\mathbf{q}} \cap \ell_{K_n,d}\big((\mathbb{R}^d)^n\big).$$

PROOF. (1) follows because $\mathbf{q}$ is globally rigid when regarded as a framework for K_n. For (2), note that if we have a framework $\tilde{\mathbf{q}}$ for $G = ([n], E)$ in $\mathbb{R}^d$, then

$$\begin{aligned} \tilde{\mathbf{q}} \text{ has the same edge lengths as } \mathbf{q} &\iff \ell_{G,d}(\tilde{\mathbf{q}}) = \ell_{G,d}(\mathbf{q}) \\ &\iff \pi_E(\ell_{K_n,d}(\tilde{\mathbf{q}})) = \ell_{G,d}(\mathbf{q}) \\ &\iff \ell_{K_n,d}(\tilde{\mathbf{q}}) \in \pi_E^{-1}(\ell_{G,d}(\mathbf{q})) = L_{\mathbf{q}}. \end{aligned}$$

Combining this with (1), we see that the desired bijection is given by

$$\text{congruence class of } \tilde{\mathbf{q}} \longmapsto \ell_{K_n,d}(\tilde{\mathbf{q}}).$$

□

This lemma implies that counting congruence classes reduces to counting points in the intersection of the affine subspace $L_{\mathbf{q}}$ and the image $\ell_{K_n,d}\big((\mathbb{R}^d)^n\big)$. However,

- if we replace the image with its Zariski closure over $\mathbb{C}$, and
- if the affine subspace has complementary dimension,

then the intersection would compute the *degree* of Zariski closure. In the projective setting, this would be the degree of the Cayley-Menger variety, which we know from Theorem 4.30. This is intriguing. Is there a theorem lurking here?

Generic Minimally Rigid Frameworks. To flesh out the ideas in the previous paragraph, we first note that the homogeneous polynomials defining the Cayley-Menger variety $CM_{n,d,\mathbb{C}} \subseteq \mathbb{P}^{\binom{n}{2}-1}$ also define its affine cone

$$^aCM_{n,d,\mathbb{C}} \subseteq \mathbb{C}^{\binom{n}{2}},$$

which is the Zariski closure of $\ell_{K_n,d}\big((\mathbb{R}^d)^n\big) \subseteq \mathbb{R}^{\binom{n}{2}} \subseteq \mathbb{C}^{\binom{n}{2}}$. Translated to the affine cone, Theorem 4.30 tells us that

$$\dim {}^aCM_{n,d,\mathbb{C}} = dn - \binom{d+1}{2}.$$

It follows that in the previous paragraph, we want to use an affine subspace of complementary dimension, i.e., of codimension $dn - \binom{d+1}{2}$. But in Lemma 4.31, $L_{\mathbf{q}}$ has codimension $m = |E|$ since it is a fiber of the projection map $\pi_E : \mathbb{R}^{\binom{n}{2}} \to \mathbb{R}^m$. Hence we will consider graphs $G = ([n], E)$ satisfying

$$m = |E| = dn - \binom{d+1}{2}.$$

For a framework $\mathbf{q}$ in $\mathbb{R}^d$ for G, we make three further assumptions:

- $\mathbf{q} = (\mathbf{q}_1, \ldots, \mathbf{q}_n)$ affinely spans $\mathbb{R}^d$.
- $\mathbf{q}$ is locally rigid for G.
- The edge lengths $\ell_{G,d}$ are generic in $\mathbb{R}^m$. Below we explain what this means.

Before stating our main theorem, we explore some consequences of these hypotheses. In Theorem 4.19, we noted that infinitesimal rigidity implies local rigidity. While the converse can fail, it holds generically, as we now explain.

THEOREM 4.32. *When the framework* $\mathbf{q}$ *is generic, infinitesimal rigidity is equivalent to local rigidity.*

PROOF. In [**11**, §3], Asimow and Roth prove that $\mathbf{q}$ is infinitesimally rigid for the graph G if and only if the Jacobian $J\ell_{G,d}$ has maximal rank at $\mathbf{q}$ and $\mathbf{q}$ is locally rigid for G. Then we are done since the rank condition is satisfied for generic $\mathbf{q}$. □

LEMMA 4.33. *The above hypotheses imply that* $\mathbf{q}$ *is minimally locally rigid for* G.

PROOF. If $\mathbf{q}$ is not minimally locally rigid for G, we could remove an edge to get a subgraph G' such that $\mathbf{q}$ would be locally rigid for G'. By the previous theorem, it would also be infinitesimally rigid for G'. Since G' has $|E| - 1 = m - 1$ edges and $\mathbf{q}$ affinely spans $\mathbb{R}^d$, Corollary 4.22 would imply that $m - 1 \geq dn - \binom{d+1}{2}$. This is impossible since we are also assuming $m = dn - \binom{d+1}{2}$. □

Here is a 2004 theorem of Borcea and Streinu [**36**].

THEOREM 4.34. *For G and $\mathbf{q}$ as above, the number of congruence classes of frameworks for G in $\mathbb{R}^d$ with the same edge lengths as $\mathbf{q}$ is bounded by*

$$\prod_{\ell=0}^{n-d-2} \frac{\binom{n-1+\ell}{n-d-1-\ell}}{\binom{2\ell+1}{\ell}}.$$

PROOF. Since ${}^aCM_{n,d,\mathbb{C}} \subseteq \mathbb{C}^{\binom{n}{2}}$ is the Zariski closure of the image of the map $\ell_{K_n,d} : (\mathbb{R}^d)^n \to \mathbb{R}^{\binom{n}{2}}$, (4.13) gives the commutative diagram of varieties:

$$\begin{array}{ccc} (\mathbb{C}^d)^n & \xrightarrow{\ell_{K_n,d}} & {}^aCM_{n,d,\mathbb{C}} \\ & \searrow^{\ell_{G,d}} & \downarrow{\scriptstyle \phi_E} \\ & & \mathbb{C}^m \end{array} \tag{4.14}$$

where ϕ_E is the restriction of the projection $\mathbb{C}^{\binom{n}{2}} \to \mathbb{C}^m$ to ${}^aCM_{n,d,\mathbb{C}}$. Our hypotheses and Theorem 4.21 imply that the rigidity matrix $R_{G,\mathbf{q}}$ has rank $dn - \binom{d+1}{2} = m$. By (4.7), $R_{G,\mathbf{q}}$ is the Jacobian of $\ell_{G,d}$ at $\mathbf{q}$ up to a factor of $\frac{1}{2}$. It follows that $\ell_{G,d}$ is dominating, which implies that ϕ_E in (4.14) is also dominating. Since

$$\dim {}^aCM_{n,d,\mathbb{C}} = dn - \binom{d+1}{2} = m,$$

we conclude that $\phi_E : {}^aCM_{n,d,\mathbb{C}} \to \mathbb{C}^m$ has finite degree. Hence there is an integer $D > 0$ such that a generic fiber of ϕ_E has D points. We assume that $\mathbf{q}$ is generic in this sense, i.e., $\phi_E^{-1}(\mathbf{q})$ has D points.

By Lemma 4.31 and the above diagram, we obtain

$$\left\{\begin{array}{l}\text{congruence classes of frameworks for } G \\ \text{in } \mathbb{R}^d \text{ with the same edge lengths as } \mathbf{q}\end{array}\right\} \simeq L_{\mathbf{q}} \cap \ell_{K_n,d}\big((\mathbb{R}^d)^n\big) \subseteq \phi_E^{-1}(\mathbf{q}).$$

When $\mathbf{q}$ is generic, there are at most D congruence classes of frameworks for G.

To finish the proof, we need to show that D is the product in the statement of the theorem. Via Theorem 4.30, this reduces to proving that $CM_{n,d,\mathbb{C}} \subseteq \mathbb{P}^{\binom{n}{2}-1}$ has degree D. To see why, note that $\phi_E : {}^aCM_{n,d,\mathbb{C}} \to \mathbb{C}^m$ is a projection of generic degree D between varieties of dimension m. This induces a rational map

$$\Phi_E : CM_{n,d,\mathbb{C}} \dashrightarrow \mathbb{P}^{m-1} \tag{4.15}$$

of generic degree D between projective varieties of dimension $m-1$, where Φ_E is the restriction of the projection map

$$\Pi_E : \mathbb{P}^{\binom{n}{2}-1} \dashrightarrow \mathbb{P}^{m-1}$$

induced by $\pi_E : \mathbb{C}^{\binom{n}{2}-1} \to \mathbb{C}^m$. It follows that for $\mathbf{q}$ generic

$$\Phi_E^{-1}(\mathbf{q}) = \Pi_E^{-1}(\mathbf{q}) \cap CM_{n,d,\mathbb{C}} \tag{4.16}$$

has D points. We will show that for generic $\mathbf{q}$, the intersection has $\deg(CM_{n,d,\mathbb{C}})$ points. The key idea is to consider three subvarieties of $CM_{n,d,\mathbb{C}}$:

- Points where Φ_E is not defined (the base locus).
- Points where $CM_{n,d,\mathbb{C}}$ is not smooth (the singular locus).
- Points where Φ_E ramifies (the ramification locus).

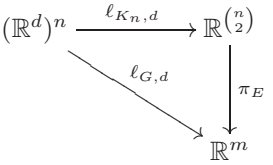
These form a proper subvariety $W \subseteq CM_{n,d,\mathbb{C}}$, hence the same is true for its image under Φ_E since Φ_E is a dominating rational map between varieties of the same dimension. Then $\mathbf{q} \in (\mathbb{R}^d)^n$ is generic when $\ell_{G,d}(\mathbf{q})$ avoids the the image of W.

Since we avoid the base locus, (4.16) can be written

$$\Phi_E^{-1}(\mathbf{q}) = \overline{\Pi_E^{-1}(\mathbf{q})} \cap CM_{n,d,\mathbb{C}},$$

where $\overline{\Pi_E^{-1}(\mathbf{q})}$ is a linear space of codimension $m-1$. Since the intersection is finite, Bézout's Theorem tells us that the intersection consists of $\deg(CM_{n,d,\mathbb{C}})$ points, counted with multiplicity. But our choice of $\mathbf{q}$ guarantees that all points of the intersection occur at smooth points of $CM_{n,d,\mathbb{C}}$ where Φ_E is unramified. This means that all of the multiplicities are 1, so that intersection has $\deg(CM_{n,d,\mathbb{C}})$ points. But it also has D points, which gives the desired equality. □

Recent Developments. The proof of Theorem 4.34 uses the affine Cayley-Menger variety ${}^aCM_{n,d,\mathbb{C}} \subseteq \mathbb{C}^{\binom{n}{2}}$ and the diagram (4.14), both of which have their origins in the commutative diagram



In contrast, the two recent papers

- [**66**] by Capco, Gallet, Grasegger, Koutschan, Lubbes and Schicho (2017)
- [**16**] by Bartzos, Emiris, Legerský and Tsigaridas (2018)

take a different point of view. They assume $d = 2$ or 3 and completely ignore K_n, focusing instead on the configuration varieties

$$C_{G,d,\ell} = \ell_{G,d}^{-1}(\ell) \subseteq (\mathbb{R}^d)^n$$

defined in (4.6). The group $SE(d)$ acts on $C_{G,d,\ell}$, and [**66**] and [**16**] find a way (different in each paper) to reduce to an explicit system of polynomial equations with finitely many solutions over $\mathbb{R}$ and $\mathbb{C}$. The number of solutions over $\mathbb{R}$ gives the number of frameworks for G in $\mathbb{R}^d$ with fixed edge lengths up to $SE(d)$ equivalence, while the number over $\mathbb{C}$ is easier to deal with (both practically and theoretically) and gives an upper bound for the number of frameworks that is much better than the bound given in Theorem 4.34. These papers use a variety of tools, including tropical geometry, bigraphs and recursion in [**66**] and mixed volumes and homotopy continuation in [**16**]. We expect further progress as people continue to study these wonderful problems.

Final Comments. This section introduced various flavors of rigidity and explored the connections to algebraic geometry. However, to unlock the full power of rigidity theory, we need to bring combinatorics into the picture. This is our next task.

4.2. Combinatorics of Rigidity

The combinatorical aspects of rigidity are best understood using the language of *matroids*. A key player is the *rigidity matroid*, which has close relations to both the rigidity matrix $R_{K_n,d}$ and the Cayley-Menger variety $CM_{n,d,\mathbb{C}}$. We will

explore some properties of this matroid and give a classic result about minimally locally rigid frameworks in the plane due to Laman and Pollaczek-Geiringer. Then algebraic statistics will make a surprise appearance.

Matroids. We begin with two examples of matroids.

EXAMPLE 4.35. Suppose we have a vector space V over a field k and a finite subset $S = \{v_1, \ldots, v_\ell\} \subseteq V$. A basic idea in linear algebra is *linear independence*, which leads to the following collection of subsets of S:

$$\mathcal{I} = \{T \subseteq S \mid T \text{ is linearly independent over } k\}.$$

The collection $\mathcal{I}$ has three interesting properties:

- $\emptyset \in \mathcal{I}$.
- $U \subseteq T \in \mathcal{I}$ implies $U \in \mathcal{I}$.
- $U, T \in \mathcal{I}$ with $|U| < |T|$ implies $\exists\, v_i \in T \setminus U$ such that $U \cup \{v_i\} \in \mathcal{I}$.

The first two properties are obvious, while the third, called the *exchange property*, takes a little thought. ◁▷

EXAMPLE 4.36. Suppose we have a field extension $k \subseteq K$ and a finite subset $S = \{\alpha_1, \ldots, \alpha_\ell\} \subseteq K$. A basic idea in field theory is *algebraic independence*, which leads to the following collection of subsets of S:

$$\mathcal{I} = \{T \subseteq S \mid T \text{ is algebraically independent over } k\}.$$

The collection $\mathcal{I}$ has three interesting properties:

- $\emptyset \in \mathcal{I}$.
- $U \subseteq T \in \mathcal{I}$ implies $U \in \mathcal{I}$.
- $U, T \in \mathcal{I}$ with $|U| < |T|$ implies $\exists\, \alpha_i \in T \setminus U$ such that $U \cup \{\alpha_i\} \in \mathcal{I}$.

The first two properties are obvious, while the third, called the *exchange property*, takes a little thought. ◁▷

The parallel language used in these examples is deliberate. In general, a *matroid* is a collection $\mathcal{I}$ of subsets of a finite set S that satisfies the properties in the examples. We call S the *ground set*, and $E \subseteq S$ is *independent* when $E \in \mathcal{I}$. Example 4.35 is called a *linear matroid*, while Example 4.36 is not surprisingly called an *algebraic matroid*.

A nice introduction to matroids (with an emphasis on algebraic matroids) can be found in the expository paper [**297**]. A more complete treatment of matroid theory appears in the book [**317**, Chapter 3]. All matroids considered here will be either linear or algebraic (sometimes both, up to isomorphism).

The Rigidity Matroid, Definition 1. Given a framework $\mathbf{q} = (\mathbf{q}_1, \ldots, \mathbf{q}_n) \in (\mathbb{R}^d)^n$ for the complete graph K_n in $\mathbb{R}^d$, the rigidity matrix $R_{K_n,\mathbf{q}}$ from Section 4.1 is the matrix that represents the linear map

$$(\mathbf{q}'_1, \ldots \mathbf{q}'_n) \in (\mathbb{R}^d)^n \longmapsto \big((\mathbf{q}_i - \mathbf{q}_j) \cdot (\mathbf{q}'_i - \mathbf{q}'_j)\big)_{ij \in \binom{[n]}{2}} \in \mathbb{R}^{\binom{n}{2}},$$

where $\binom{[n]}{2}$ is edge set of K_n defined in (4.12). To simplify notation, we write $R_{\mathbf{q}}$ instead of $R_{K_n,\mathbf{q}}$. This leads to our first definition of the rigidity matroid.

DEFINITION 4.37. Pick a point $\mathbf{q} \in (\mathbb{R}^d)^n$ whose coordinates are algebraically independent over $\mathbb{Q}$. Then *rigidity matroid* $\mathcal{R}_d(n)$ is the linear matroid whose independent sets are the linearly independent rows of the rigidity matrix $R_{\mathbf{q}}$.

If $\tilde{\mathbf{q}}$ is another such set of points, there is an automorphism $\sigma : \mathbb{R} \to \mathbb{R}$ with $\sigma(\mathbf{q}) = \tilde{\mathbf{q}}$. It follows easily that $R_{\mathbf{q}}$ and $R_{\tilde{\mathbf{q}}}$ give the same matroid. Thus $\mathcal{R}_d(n)$ is the linear matroid of the rows of the rigidity matrix of a framework for K_n in $\mathbb{R}^d$ that is as generic as possible.

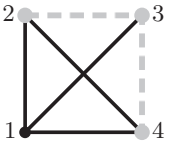
Since the rows of the rigidity matrix $R_{\mathbf{q}}$ are indexed by $\binom{[n]}{2}$, we can use $\binom{[n]}{2}$ as the ground set for the rigidity matroid $\mathcal{R}_d(n)$. Then a subset $E \subseteq \binom{[n]}{2}$ picks out a collection of rows in $R_{\mathbf{q}}$ and simultaneously determines a subgraph of K_n. This is our next topic.

Edge Subgraphs of K_n. Take a subset $E \subseteq \binom{[n]}{2}$ and let

$$V(E) = \{\text{vertices of } K_n \text{ incident to an edge in } E\} \subseteq [n].$$

We call $G(E) = (V(E), E)$ an *edge subgraph* of K_n since it is determined by its set of edges E.

EXAMPLE 4.38. If $n = 4$ and $E = \{23, 34\} \subseteq \binom{[4]}{2}$, then $V(E) = \{2, 3, 4\}$.



The edge subgraph $G(E)$ is in shown in gray with dashed edges. ◁▷

Given $E \subseteq \binom{[n]}{2}$ and a generic framework $\mathbf{q} \in (\mathbb{R}^d)^n$, we let $\mathbf{q}_E$ be the elements of $\mathbf{q}$ indexed by $V(E)$. It follows that $\mathbf{q}_E$ is a generic framework for $G(E)$ in $\mathbb{R}^d$. This framework has a rigidity matrix $R_{G(E),\mathbf{q}_E}$ that relates to $R_{\mathbf{q}}$ as follows (we omit the straightforward proof).

LEMMA 4.39. *After suitably rearranging the columns of $R_{\mathbf{q}}$, we have*

$$\big[\text{rows of } R_{\mathbf{q}} \text{ indexed by } E\big] = \big[R_{G(E),\mathbf{q}_E} \mid 0\ \big].$$

Thus $R_{\mathbf{q}}$ contains the rigidity matrices of all edge subgraphs $G(E)$ of K_n. But which of these correspond to independent sets in $\mathcal{R}_d(n)$?

Two Theorems. We next state two basic results about the independent sets $E \in \mathcal{R}_d(n)$. We refer the reader to [**317**, Theorems 3.11.2 and 3.11.3] for proofs.

THEOREM 4.40. *If E is independent in $\mathcal{R}_d(n)$, then every subset $F \subseteq E$ with $|V(F)| \geq d$ satisfies $|F| \leq d|V(F)| - \binom{d+1}{2}$.*

THEOREM 4.41. *Consider $E \subseteq \binom{[n]}{2}$. Then:*

(1) *If $|V(E)| \leq d+1$, then E is independent in $\mathcal{R}_d(n)$.*
(2) *If $|V(E)| \geq d+1$, then any two of the following imply the third:*
   - *E is independent in $\mathcal{R}_d(n)$.*
   - *E is locally rigid.*
   - *$|E| = d|V(E)| - \binom{d+1}{2}$.*

To understand the difference between $|V(E)| \leq d+1$ and $|V(E)| > d+1$, note that since $\mathbf{q} = (\mathbf{q}_1, \ldots, \mathbf{q}_n)$ is generic, the convex hull of $\mathbf{q}$ in $\mathbb{R}^d$ is a simplex when $|V(E)| \leq d+1$, and independence is clear. But things can be more complicated

when $|V(E)| \geq d+1$. Here, Corollary 4.22 and Theorem 4.32 tell us:

- Independence implies $|E| \leq d|V(E)| - \binom{d+1}{2}$.
- Local rigidity implies $|E| \geq d|V(E)| - \binom{d+1}{2}$.

This shows that the the first two bullets in Theorem 4.41 imply the last. The pleasant surprise is that any two imply the third. There are many ideas involved here, but they interact very nicely.

The Laman–Pollaczek-Geiringer Theorem. A classic result about the rigidity of frameworks in $\mathbb{R}^2$ was discovered, forgotten, and rediscovered:

- In 1927, Hilda Pollaczek-Geiringer published *Über die Gliederung ebener Fachwerke* [**294**] in Zeitschrift für Angewandte Mathematik und Mechanik.
- In 1970, Gerard Laman published *On graphs and rigidity of plane skeletal structures* [**250**] in the Journal of Engineering Mathematics.

Laman was unaware of Pollaczek-Geiringer's earlier paper. Here is the main result proved in their papers.

THEOREM 4.42. *If $G = (V, E)$ with $|V| \geq 2$, then a generic framework in $\mathbb{R}^2$ for G is minimally locally rigid if and only if*

$$|E| = 2|V| - 3 \text{ and } |F| \leq 2|V(F)| - 3 \text{ for all } \emptyset \neq F \subseteq E. \tag{4.17}$$

PROOF. The theorem is trivial when $|V| = 2$. We will prove one direction when $|V| \geq 3$. Suppose that $\mathbf{q}$ is generic and minimally locally rigid for G in $\mathbb{R}^2$. By Theorem 4.32, it is infinitesimally rigid, so that by Theorem 4.21,

$$|E| \geq \operatorname{rank}(R_{G,\mathbf{q}}) = 2|V| - 3.$$

Rows of the rigidity matrix are indexed by edges, so $2|V| - 3$ linearly independent rows give the rigidity matrix of a subgraph of G. By Theorem 4.21, this subgraph is infinitesimally rigid, hence locally rigid. Since $\mathbf{q}$ is minimal for G, the subgraph must equal G, which implies $|E| = 2|V| - 3$. It follows that E is independent, and then subsets $\emptyset \neq F \subseteq E$ satisfy (4.17) by Theorem 4.40.

For the other direction, one can give a matroidal proof [**317**, Theorem 4.2.1]. Another approach uses a recursive argument based on *Henneberg moves*. See, for example, [**66**, Appendix A]. □

For many years, Theorem 4.42 was known as "Laman's Theorem," but now "Laman–Pollaczek-Geiringer Theorem" is more common. As noted in [**16**], James Clerk Maxwell knew part of the theorem in 1864, when he wrote:

> A frame of s points in a plane requires in general $2s-3$ connecting lines to render it stiff. [**268**]

The Rigidity Matroid, Definition 2. We next present $\mathcal{R}_d(n)$ as a algebraic matroid using the affine Cayley-Menger variety ${}^aCM_{n,d,\mathbb{C}} \subseteq \mathbb{C}^{\binom{n}{2}}$. Recall that this is the Zariski closure of the image of the length-squared map $\ell_{K_n,d} : (\mathbb{R}^d)^n \to \mathbb{R}^{\binom{n}{2}}$.

Using coordinates $\ell_{ij},\ ij \in \binom{[n]}{2}$, on $\mathbb{C}^{\binom{n}{2}}$, ${}^aCM_{n,d,\mathbb{C}}$ is defined by a prime ideal $P_{n,d} \subseteq R = \mathbb{C}\big[\ell_{ij} \mid ij \in \binom{[n]}{2}\big]$. Its field of rational functions is $R/P_{n,d} \hookrightarrow \mathbb{K}^{n,d}$. Thus $\mathbb{K}^{n,d}$ is generated over $\mathbb{C}$ by the cosets of the ℓ_{ij} in $R/P_{n,d}$.

We now give a second definition of $\mathcal{R}_d(n)$.

DEFINITION 4.43. The *rigidity matroid* $\mathcal{R}_d(n)$ is the algebraic matroid whose independent sets are subsets $E \subseteq \binom{[n]}{2}$ such that either of the following equivalent conditions holds:

(1) The ℓ_{ij}, $ij \in E$, are algebraically independent over $\mathbb{C}$ as rational functions on ${}^aCM_{n,d,\mathbb{C}}$, or
(2) $P_{n,d} \cap \mathbb{C}\big[\ell_{ij} \mid ij \in E\big] = \{0\}$.

The equivalence follows because (2) gives $\mathbb{C}\big[\ell_{ij} \mid ij \in E\big] \hookrightarrow R/P_{n,d} \hookrightarrow \mathbb{K}^{n,d}$. As a consequence, we can compute the independent sets of $\mathcal{R}_d(n)$ by elimination theory (though this is wildly impractical except in very simple examples).

Equivalence of Definitions 1 and 2. Here is the theorem we expect.

THEOREM 4.44. *The linear matroid definition of $\mathcal{R}_d(n)$ via the rigidity matrix $R_{\mathbf{q}}$ is equivalent to the algebraic matroid definition of $\mathcal{R}_d(n)$ via the Cayley-Menger variety ${}^aCM_{n,d,\mathbb{C}}$.*

PROOF. The construction of ${}^aCM_{n,d,\mathbb{C}}$ implies that the complex version of the length-squared map

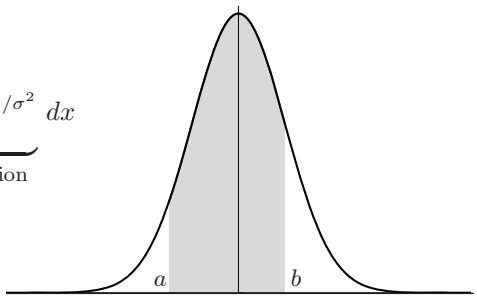
$$\ell_{K_n,d} : (\mathbb{C}^d)^n \longrightarrow {}^aCM_{n,d,\mathbb{C}} \subseteq \mathbb{C}^{\binom{n}{2}}$$

is dominating. At a generic point $\mathbf{q} \in (\mathbb{C}^d)^n$, the matrix of $\frac{1}{2}J\ell_{K_n,d}(\mathbf{q})$ is the rigidity matrix $R_{\mathbf{q}}$ used in the linear matroid definition of $\mathcal{R}_d(n)$. From here, the desired equivalence follows without difficulty. □

Earlier, we used the linear matroid interpretation of $\mathcal{R}_d(n)$ in our discussion of the Laman–Pollaczek-Geiringer Theorem. We next give an application of the algebraic matroid interpretation of $\mathcal{R}_d(n)$ to algebraic statistics, following Gross and Sullivant [**180**]. But first, we need to review some basic statistics.

Gaussian Normal Distributions. The statistical models encountered so far in this book have been discrete models. We now turn to a class of continuous models based on the *Gaussian normal distribution* $\mathcal{N}(0, \sigma^2)$, where the probability that a random variable $X \in \mathbb{R}$ lies the interval $[a, b]$ is given by the classic formula

$$P(a \le X \le b) = \int_a^b \underbrace{\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}x^2/\sigma^2}}_{\text{density function}}\, dx$$



We generalize this to n dimensions as follows:

- $X = (X_1, \ldots, X_n)$ is a random variable in $\mathbb{R}^n$ distributed according to $\mathcal{N}(0, \Sigma)$, the *multivariable Gaussian normal distribution* with *mean* $0 \in \mathbb{R}^n$ and *covariance matrix* Σ.
- Σ is a $n \times n$ symmetric positive definite matrix (when $n = 1, \Sigma = (\sigma^2)$).

- The *density function* for $x \in \mathbb{R}^n$ is

$$f(x) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} \exp\bigl(-\tfrac{1}{2} x^t \Sigma^{-1} x\bigr).$$

The covariance matrix Σ is the *parameter* that defines the model, and its inverse Σ^{-1} is usually called the *concentration matrix*.

Gaussian Graphical Models. Suppose that we have a graph $G = (V, E)$ with $V = [n]$ as usual.

DEFINITION 4.45. A *Gaussian graphical model with graph* G is a Gaussian normal distribution $\mathcal{N}(0, \Sigma)$ whose concentration matrix Σ^{-1} satisfies

$$(\Sigma^{-1})_{ij} = 0 \quad \text{for } i \neq j \text{ and } ij \notin E. \tag{4.18}$$

The vanishing condition says that when $i \neq j$ and $ij \notin E$, the random variables X_i and X_j are conditionally independent relative to the remaining variables.

Maximum Likelihood Estimates. In the discrete statistical models considered in Sections 2.2 and 3.1, the *maximum likelihood estimate* played an important role. Recall that the MLE is (roughly speaking) the model that best explains the data.

To apply this idea to a Gaussian normal distribution $\mathcal{N}(0, \Sigma)$, suppose we have observations $X^{(1)}, \ldots, X^{(d)} \in \mathbb{R}^n$ with mean 0. Then the maximum likelihood estimate is the choice $\hat{\Sigma}$ of the parameter Σ that maximizes the *log-likelihood function* for the given data. We explain this as follows. The data gives the *sample covariance matrix*

$$S = \frac{1}{d} \sum_{i=1}^{d} X^{(i)} (X^{(i)})^t. \tag{4.19}$$

This is a positive semidefinite $n \times n$ symmetric matrix of rank $\leq d$ since it is the sum of d rank 1 matrices $X^{(i)}(X^{(i)})^t$. In terms of S, the log-likelihood is

$$-\log\det(\Sigma) - \operatorname{tr}(S\Sigma^{-1}).$$

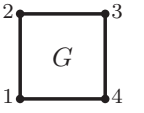
Now suppose that our model is unconstrained, i.e., Σ is allowed to be any $n \times n$ positive definite matrix. If in addition $d \geq n$, i.e., if the number of observations is at least the dimension, then with probability 1, the sample covariance matrix S has rank n and hence is positive definite. In this situation, one can show that $\hat{\Sigma} = S$ maximizes the log-likelihood function. Thus, in the unconstrained setting, the MLE exists with probability 1 when $d \geq n$.

However, when we are dealing with a Gaussian graphical model for $G = (V, E)$, the covariance matrix Σ is constrained to satisfy $(\Sigma^{-1})_{ij} = 0$ when $i \neq j$ and $ij \notin E$. Hence the above analysis no longer applies.

The MLE of a Gaussian Graphical Model. Given $G = (V, E)$ and observations $X^{(1)}, \ldots, X^{(d)} \in \mathbb{R}^n$, the restriction on Σ^{-1} means that in the sample covariance matrix S from (4.19), only the entries S_{ij} with $i = j$ or $ij \in E$ have any meaning. This leads to the *partial matrix*

$$S_G = (S_{ij})_{i=j \text{ or } ij \in E}.$$

EXAMPLE 4.46. Here is a graph G and possible sample covariance matrix S:



$$S = \begin{bmatrix} 1 & .9 & ? & -.9 \\ .0 & 1 & .9 & ? \\ ? & .9 & 1 & .9 \\ -.9 & ? & .9 & 1 \end{bmatrix}$$

The question marks indicate numbers that mean nothing for the model. They correspond to the diagonals of the square, which are not edges of the graph. We obtain S_G by removing these entries. ◁▷

Using partial matrices, we can describe when the MLE exists as follows.

THEOREM 4.47. *In the Gaussian graphical model for $G = ([n], E)$ and data with sample covariance matrix S, the* MLE *exists if and only if there is a positive definite Σ satisfying $\Sigma_G = S_G$.*

We learned about this standard result from Uhler's 2012 paper *Geometry of maximum likelihood estimation in Gaussian graphical models* [**349**, Theorem 2.1]. The matrix Σ in the statement of the theorem is not unique—to get the MLE $\hat{\Sigma}$, let $\hat{\Sigma}$ be the Σ with maximal determinant that satisfies $\Sigma_G = S_G$. One can prove that this maximum exists, is positive definite, and satisfies (4.18).

When $d \geq n$, S is positive definite with probability 1 and we can take $\Sigma = S$ in Theorem 4.47 in this case. Hence we have proved the following corollary.

COROLLARY 4.48. *In the Gaussian graphical model for $G = ([n], E)$, the* MLE *exists with probability 1 for d observations when $d \geq n$.*

Gaussian Graphical Models and Rigidity. The theory of Gaussian graphical models is very nice, but how does it relate to rigidity? One unexpected connection is the following theorem of Gross and Sullivant from 2018 [**180**, Theorem 1.2].

THEOREM 4.49. *In the Gaussian graphical model for $G = ([n], E)$, the* MLE *exists with probability 1 for d observations when E is an independent set in the rigidity matroid $\mathcal{R}_{d-1}(n)$.*

This theorem is really a result about matroids. In a Gaussian graphical model, we have d observations in n-dimensional space, while $\mathcal{R}_{d-1}(n)$ is determined by the rigidity matrix of a framework with n points in $(d-1)$-dimensional space. In what follows, we first discuss a theorem of Uhler [**349**] that relates existence of the MLE to an algebraic matroid built from symmetric matrices and then explain how Gross and Sullivant relate this matroid to $\mathcal{R}_{d-1}(n)$.

The Symmetric Matroid and Uhler's Theorem. By (4.19), d observations $X^{(1)}, \ldots, X^{(d)} \in \mathbb{R}^n$ give the $n \times n$ sample covariance matrix

$$S = \frac{1}{d} \sum_{i=1}^{d} X^{(i)} (X^{(i)})^t,$$

which is symmetric positive semidefinite of rank $\leq d$. This gives an affine variety over $k = \mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$ as follows. Introduce $\binom{n+1}{2}$ variables x_{ij} for $1 \leq i \leq j \leq n$ on the affine space $k^{\binom{n+1}{2}}$, and let

$$I_{n,d,k} \subseteq k[x_{ij} \mid 1 \leq i \leq j \leq n]$$

be the ideal of generated by all $(d+1)\times(d+1)$ minors of the $n \times n$ symmetric matrix built from the x_{ij}:

$$\begin{bmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1n} \\ x_{12} & x_{22} & x_{23} & \cdots & x_{2n} \\ x_{13} & x_{23} & x_{33} & \cdots & x_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{1n} & x_{2n} & x_{3n} & \cdots & x_{nn} \end{bmatrix}.$$

It is well-known that $I_{n,d,k}$ is a prime ideal.

We will consistently abuse notation and regard $n \times n$ symmetric matrices as elements of $k^{\binom{n+1}{2}}$. Under this identification, the affine variety

$$\mathrm{Sym}_{n,d,k} = \mathbf{V}(I_{n,d,k}) \subseteq k^{\binom{n+1}{2}}$$

consists of all $n \times n$ symmetric matrices of rank $\leq d$ with entries in k. In practice, $k = \mathbb{Q}$ for symbolic computation, $k = \mathbb{R}$ for many applications, and $k = \mathbb{C}$ for algebraic geometry.

The variables x_{ij} give rational functions on the irreducible variety $\mathrm{Sym}_{d,n,k}$. This is exactly the information needed to define an algebraic matroid.

Definition 4.50. The *symmetric matroid* $\mathcal{S}_d(n)$ is the algebraic matroid of the variety $\mathrm{Sym}_{d,n,k}$ relative to the rational functions x_{ij}.

This matroid is independent of k, which justifies the notation $\mathcal{S}_d(n)$. In 2012, Uhler proved the following [**349**, Theorem 3.3].

Theorem 4.51. *For $G = (V, E)$ with $V = [n]$, the MLE exists with probability 1 for d observations in a Gaussian graphical model $\mathcal{N}(0, \Sigma)$ for G when either of the following equivalent conditions are satisfied*:

(1) $I_{n,d,k} \cap k[x_{ii}, x_{ij} \mid i \in V, ij \in E] = \{0\}$, *or*
(2) $\{x_{ii}, x_{ij} \mid i \in V, ij \in E\}$ *is an independent set in the symmetric matroid* $\mathcal{S}_d(n)$.

Proof. The equivalence of (1) and (2) follows because (1) gives the inclusion

$$k[x_{ii}, x_{ij} \mid i \in V, ij \in E] \hookrightarrow k[x_{ii}, x_{ij}]/I_{n,d,k}.$$

The variables x_{ii} for $i \in V$ and x_{ij} for $ij \in E$ correspond precisely to the entries of the partial matrix S_G.

When $k = \mathbb{R}$, these variables give coordinates on $\mathbb{R}^{V+E}$, and the map sending a symmetric matrix S to its partial matrix S_G is represented by the projection

$$\pi_G : \mathbb{R}^{\binom{n+1}{2}} \longrightarrow \mathbb{R}^{V+E}.$$

We can interpret Theorem 4.49 in terms of π_G as follows. Let $\mathrm{PD}_n \subseteq \mathbb{R}^{\binom{n+1}{2}}$ be the cone of $n \times n$ symmetric positive definite matrices. Its image under π_G is the *cone of sufficient statistics*

$$\mathcal{C}_G = \pi_G(\mathrm{PD}_n) \subseteq \mathbb{R}^{V+E}.$$

For a sample covariance matrix S, the condition in Theorem 4.49 that $S_G = \Sigma_G$ for some $\Sigma \in \mathrm{PD}_n$ is equivalent to saying

$$S_G \in \mathcal{C}_G.$$

We now prove the theorem. Since the MLE exists with probability 1 when $d \geq n$ (Corollary 4.48), we can assume $d < n$. Our proof will use standard facts about semialgebraic sets. A good reference is [**93**].

When S comes from d observations, it lies in the cone $\mathrm{PSD}_{n,d}$ of positive semidefinite matrices of rank $\leq d$, so that

$$S_G \in \mathcal{V} = \pi_G(\mathrm{PSD}_{n,d}) \subseteq \mathbb{R}^{V+E}.$$

Since $\mathrm{PSD}_{n,d}$ is semialgebraic (defined by polynomial equalities and inequalities), the same is true for $\mathcal{V}$ by the Tarski-Seidenberg Theorem [**93**, Corollary 2.4].

Let's study how $\mathcal{V}$ interacts with $\mathcal{C}_G$. Since $\mathrm{PSD}_{n,d} \subseteq \overline{\mathrm{PD}}_n$, we have

$$\mathcal{V} \subset \pi_G(\overline{\mathrm{PD}}_n) \subseteq \overline{\pi_G(\mathrm{PD}_n)} = \overline{\mathcal{C}}_G = \mathcal{C}_G \cup \partial\mathcal{C}_G.$$

Here, $\overline{(\cdots)}$ is topological closure and $\partial(\cdots)$ is topological boundary. Hence

$$\mathcal{V} = (\mathcal{V} \cap \mathcal{C}_G) \cup (\mathcal{V} \cap \partial\mathcal{C}_G). \tag{4.20}$$

This union is disjoint since $\mathcal{C}_G$ is open (being a projection of the open cone PD_n).

Our hypothesis involves the ideal $I_{n,d,\mathbb{R}}$ defining $\mathrm{Sym}_{n,d,\mathbb{R}}$. The relation between this variety and $\mathrm{PSD}_{n,d}$ is easy to describe: we have an inclusion $\mathrm{PSD}_{n,d} \subseteq \mathrm{Sym}_{n,d,\mathbb{R}}$, and [**201**, Proposition 5.1.1] easily implies that

$$\mathrm{PSD}_{n,d} \text{ is Zariski dense in } \mathrm{Sym}_{n,d,\mathbb{R}}. \tag{4.21}$$

This is what happens in $\mathbb{R}^{\binom{n+1}{2}}$. But what about in $\mathbb{R}^{V+E}$? We claim that

$$\mathcal{V} \text{ is Zariski dense in } \mathbb{R}^{V+E}. \tag{4.22}$$

To see why, suppose that $f \in \mathbb{R}[x_{ii}, x_{ij} \mid i \in V, ij \in E]$ vanishes on $\mathcal{V}$. Then f, regarded as a polynomial function on $\mathbb{R}^{\binom{n+1}{2}}$, vanishes on $\mathrm{PSD}_{n,d}$, which by (4.21) implies that f vanishes on $\mathrm{Sym}_{n,d,\mathbb{R}}$. But $\mathrm{Sym}_{n,d,\mathbb{R}}$ is Zariski dense in $\mathrm{Sym}_{n,d,\mathbb{C}}$, so the Nullstellensatz implies that $f \in I_{n,d,\mathbb{C}}$ since $I_{n,d,\mathbb{C}}$ is prime. Thus

$$f \in I_{n,d,\mathbb{C}} \cap \mathbb{C}[x_{ii}, x_{ij} \mid i \in V, ij \in E] = \{0\},$$

where the last equality is hypothesis (1) of the theorem with $k = \mathbb{C}$. This completes the proof of (4.22).

Let $N = |V| + |E|$ be the dimension of $\mathbb{R}^{V+E}$. Since a semialgebraic set and its Zariski closure have the same dimension [**93**, Theorem 3.20], (4.22) implies that $\dim \mathcal{V} = N$. By [**93**, Proposition 3.15], we conclude that the topological interior $\mathcal{V}^\circ$ is a nonempty open set of $\mathbb{R}^{V+E}$.

The final fact we need is that since $\mathcal{C}_G$ is open and semialgebraic in $\mathbb{R}^{V+E}$, its boundary $\partial\mathcal{C}_G$ is semialgebraic of dimension $< N$ by [**93**, Corollary 2.5 and Proposition 3.16].

The theorem now follows. The previous two paragraphs imply that $\mathcal{V}^\circ \subseteq \partial\mathcal{C}_G$ is impossible, so that by (4.20), $\mathcal{V}^\circ \cap \mathcal{C}_G$ is a nonempty open subset of $\mathbb{R}^{V+E}$. It follows that for Lebesgue measure on $\mathbb{R}^{V+E}$, $\mathcal{V} \cap \mathcal{C}_G$ has positive measure while $\mathcal{V} \cap \partial\mathcal{C}_G$ measure zero since $\dim \partial\mathcal{C}_G < N$. By (4.20), the partial matrix $S_G \in \mathcal{V}$ lies in $\mathcal{C}_G$ with probability 1, and the theorem is proved. □

The intersection $I_{n,d,\mathbb{C}} \cap \mathbb{C}[x_{ii}, x_{ij} \mid i \in V, ij \in E]$ is an example of an *elimination ideal*. We saw in Section 1.1 how such ideals were present at the birth of elimination

theory and are central to the modern approach [**104**, Chapter 3]. One fun aspect of the above proof is how this classical idea interacts with semialgebraic geometry.

Back to the Rigidity Matroid. The symmetric matroid $\mathcal{S}_d(n)$ is nice, but our real object of study is the rigidity matroid $\mathcal{R}_d(n)$. The relation between these matroids was discovered by Gross and Sullivant in 2014. The key idea is to *remove the diagonal*.

By definition, $\mathcal{S}_d(n)$ is the algebraic matroid of $\mathrm{Sym}_d(n)$ relative to the x_{ij} for $1 \le i \le j \le n$. One can show that

$$\mathrm{Diag} = \{x_{ii} \mid i \in [n]\}$$

is independent in $\mathcal{S}_d(n)$. The *contraction matroid* $\mathcal{S}_d(n)/\mathrm{Diag}$ has ground set x_{ij} for $1 \le i < j \le n$, and given $A \subseteq \{x_{ij} \mid 1 \le i < j \le n\}$, we have

$$A \text{ is independent in } \mathcal{S}_d(n)/\mathrm{Diag} \iff \mathrm{Diag} \cup A \text{ is independent in } \mathcal{S}_d(n). \tag{4.23}$$

Here is a key result proved by Gross and Sullivant [**180**, Theorem 2.4]. We refer to their paper for the proof.

THEOREM 4.52. *The obvious bijection* $\{x_{ij} \mid 1 \le i < j \le n\} \simeq \binom{[n]}{2}$ *induces an isomorphism of matroids* $\mathcal{S}_d(n)/\mathrm{Diag} \simeq \mathcal{R}_{d-1}(n)$.

We now prove the theorem of Gross and Sullivant stated earlier in the section as Theorem 4.49:

$$\begin{aligned} E \text{ independent in } \mathcal{R}_{d-1}(n) &\iff \{x_{ij} \mid ij \in E\} \text{ independent in } \mathcal{S}_d(n)/\mathrm{Diag} \\ &\iff \mathrm{Diag} \cup \{x_{ij} \mid ij \in E\} \text{ independent in } \mathcal{S}_d(n) \\ &\iff \text{for } G = ([n], E), \text{ the MLE exists with} \\ &\qquad \text{probability 1 for } d \text{ observations.} \end{aligned} \tag{4.24}$$

The first equivalence is Theorem 4.52, the second is (4.23), and finally the third is Theorem 4.51.

REMARK 4.53. Theorem 4.52 is not the first relation we have seen between rigidity and symmetric matrices. Recall that the proof of Theorem 4.34 used the isomorphism (4.11). At the level of affine varieties, the isomorphism becomes

$$\begin{array}{ccc} \mathrm{Sym}_{n-1,d,\mathbb{C}} & \simeq & {}^aCM_{n,d,\mathbb{C}} \\ \updownarrow & & \updownarrow \\ \mathcal{S}_d(n-1) & & \mathcal{R}_d(n) \end{array} \tag{4.25}$$

In this display, the bottom row indicates the associated algebraic matroids. At first glance, this seems incompatible with

$$\mathcal{S}_d(n)/\mathrm{Diag} \simeq \mathcal{R}_{d-1}(n)$$

from Theorem 4.52. What happens is that the isomorphism (4.25) of varieties uses the change of variables $s_{ij} = \frac{1}{2}(\ell_{i1} + \ell_{j1} - \ell_{ij})$, which doesn't preserve the ground sets of the two algebraic matroids. In other words, (4.25) does not induce an isomorphism of the associated matroids. This is a nice illustration of the importance of being able work directly at the level of matroids as in Theorem 4.52.

Maximum Likelihood Estimate of a Planar Graph. As an application of the link (4.24) between MLE and the rigidity matroid, we conclude this section with

a surprising result of Gross and Sullivant about the MLE of a Gaussian graphical model of a planar graph [**180**, Corollary 3.11].

THEOREM 4.54. *Let $G = ([n], E)$ be a planar graph. Then the MLE of the Gaussian graphical model for G exists with probability 1 for $d = 4$ observations.*

PROOF. If $d = 4 \geq n$, then we are done by Corollary 4.48. Hence we may assume $n \geq 4$. By adding edges between existing vertices, we get a *maximal planar graph* $G' = ([n], E')$ with $E \subseteq E'$. We will prove that E' is independent in the rigidity matroid $\mathcal{R}_3(n)$. Then E will be independent, which by (4.24) for $d = 4$ will imply that the MLE for G exists with probability 1 for 4 observations.

Here are some standard facts about the maximal planar graph $G' = ([n], E')$:

- When regarded as lying in the sphere, maximality implies that every face of G' is a triangle. Then the number of faces f' and number of edges $m' = |E'|$ satisfy $3f' = 2m'$.
- Combining this with Euler's formula $n - m' + f' = 2$, we see that G' satisfies $m' = 3n - 6$.
- Maximality also implies that G' is *3-connected* when $n \geq 4$, meaning that it remains connected after removing any two vertices. See [**132**, Corollary 4.4.7] for a proof.

A 1916 theorem of Steinitz (see [**370**, Theorem 4.1]) states that every 3-connected planar graph can be realized in $\mathbb{R}^3$ as the edges of a convex polytope. When applied to G', the resulting polytope has triangular faces, which allows us to assume that the vertices are generic. Theorem 4.25 implies that the resulting edge framework $\mathbf{q}$ in $\mathbb{R}^3$ for G' is locally rigid. Since it also satisfies the $m' = 3n - 6$, Theorem 4.41 implies that E' is independent in the rigidity matroid $\mathcal{R}_3(n)$, as desired. □

To see the power of this theorem, suppose that we have a Gaussian graphical model for a planar graph with 10,000 vertices. How many observations d do we need for the MLE to exist with probability 1? By Corollary 4.48, we know that $d = 10{,}000$ will suffice. If instead we use Theorem 4.54, we can let $d = 4$!

Final Comments. It can take a while to get used to the language of matroids. Yet in this section we have used matroids to link a variety of topics, including rigidity, polytopes, real algebraic geometry, planar graphs, and algebraic statistics.

4.3. Polynomial Methods and Rigidity Theory

by Jessica Sidman

In Section 4.2 we saw the Laman–Pollaczek-Geiringer (LPG) theorem, which says that in the plane *generic* rigidity has a combinatorial characterization. Three questions follow naturally:

(1) Can the LPG theorem be generalized to higher dimensions, in particular to dimension three where there are potential applications to chemistry, engineering, and architecture?
(2) The LPG theorem deals with frameworks that are given by a system of distance constraints between pairs of points. Are there similar statements for other kinds of constraint systems?

(3) What can we say about when genericity fails? In particular, can we use algebraic methods to find intuitive geometric conditions under which genericity fails?

Rigidity Theory in Higher Dimensions. Let us begin by briefly recalling the intuition behind the sparsity conditions in the LPG theorem which holds for a graph $G = (V, E)$ in the plane where

(1) $|E| = 2|V| - 3$
(2) if $V' \subseteq V$ has at least two elements and E' is the edge set of the induced graph on V', then $|E'| \leq 2|V'| - 3$.

Each vertex in the plane has two degrees of freedom, and constraints between joints can only remove relative motions. The 3-dimensional space of rigid motions of the framework as a single block can never be eliminated by adding bars between joints. Hence, if the edges in G impose independent conditions there can be at most $2|V|-3$ of them.

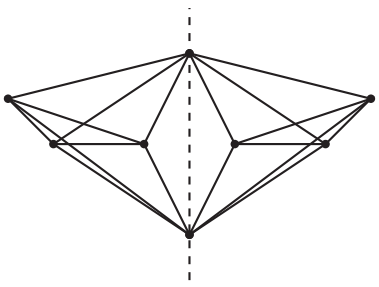
The group of rigid motions of $\mathbb{R}^d$ has dimension $\binom{d+1}{2}$, so in dimension three the space of rigid motions has dimension six. Since each vertex in a graph G has three degrees of freedom in space, the natural generalization of the conditions in the LPG theorem (also referred to as the *Maxwell counting conditions*) in $\mathbb{R}^3$ is

(1) $|E| = 3|V| - 6$
(2) if $V' \subseteq V$ has at least two elements, and E' is the edge set of the induced graph on V', then $|E'| \leq 3|V'| - 6$.

However, the famous "double banana" shows that the Maxwell counting conditions do not characterize rigidity in dimension three.

Example 4.55. The double banana depicted below is a graph with 8 vertices and $3 \cdot 8 - 6 = 18$ edges. Each "banana" is the union of two tetrahedra glued along a facet. The two bananas are glued to each other at the top and bottom vertices.

The reader may check that the Maxwell counting conditions are satisfied on all induced subgraphs. However, we may hold one banana fixed and rotate the other around the dashed line through the vertices at which the bananas are joined. This gives a nontrivial motion of the joints in one banana relative to the joints in the other one.



◁▷

The analogue of the Maxwell counting conditions fails in higher dimensions as well. For each odd dimension d there exists a graph G with $|E| = d|V| - \binom{d+1}{2}$ and $|E'| \leq d|V'| - 3\binom{d+1}{2}$ for all induced subgraphs that fails to be rigid. Examples of such graphs can be constructed as "hyperbananas" as in [**80**].

Question: Is there a combinatorial characterization of generic rigidity in dimension three? In higher dimensions?

An Introduction to Body-Based Rigidity. Though a combinatorial characterization of generic rigidity for bar-and-joint frameworks in dimension three and higher remains open, results characterizing generic rigidity exist for other types of frameworks, in particular body-based frameworks, which we introduce here.

In the 1980s a combinatorial characterization of rigidity for body-and-bar frameworks was given by Tay [**342**]. A second proof was given by White and Whiteley [**361**], who also laid out a way to analyze the combinatorial substructures of such frameworks using polynomials. To motivate the definition, consider Example 4.56.

EXAMPLE 4.56. Let A and B be two polygons in the plane. In the figure below A is the triangle and B is the square. We fix A and analyze relative motions.

If we attach a bar to both A and B, then B can rotate about its point of attachment to the bar and also translate in space.

Adding a second bar between A and B leaves a shearing motion.

Generically adding the third bar creates a rigid structure.

◁▷

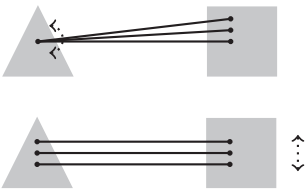
In general, we want a mathematical model that describes a situation in which we have n d-dimensional rigid bodies connected by a collection of m fixed length bars attached at universal joints. In the plane, we imagine the joints as pins about which the bodies can rotate. As Example 4.56 shows, in order to eliminate the relative motions between just two rigid bodies we will need multiple bars. Hence, it is natural to give a body-and-bar framework the underlying combinatorial structure of a multigraph. We reprise Defintion 4.3 from Section 4.1 in a form commonly found in the literature.

DEFINITION 4.57 ([**361**]). A d-dimensional *body-and-bar framework* consists of a multigraph $G = (V, E)$ and an element $(a_e, b_e) \in \mathbb{R}^d \times \mathbb{R}^d$ for each edge $e \in E$. Each vertex represents a body, and each edge e represents a bar with endpoints a_e and b_e.

In dimension 3, a famous example of such a structure is a special case of the Stewart-Gough mechanism which consists of two metal plates connected by six legs attached at ball-and-socket joints. In the full Stewart-Gough platform, the bars have adjustable length, but here we consider them fixed so that they represent fixed distances between joints. See Examples 4.4 and 4.8 from Section 4.1.

The conditions characterizing generic body-and-bar rigidity given by Tay and White and Whiteley are similar to those appearing in the LPG theorem. Again, the proof involves linearizing the problem, so the statement is for infinitesimal rigidity. Moreover, as we see in Example 4.58, any combinatorial characterization of rigidity can only hold for a sufficiently general framework associated to a given multigraph.

EXAMPLE 4.58. Three bars with sufficiently general attachment points on two rigid bodies in the plane eliminate all relative motion. However, as we see below, there are special placements of the three bars that allow internal motion of the resulting framework.



If the three bars meet at a point, then the square can rotate around the cube.

If the three bars are parallel, then the system admits a shearing motion. ◁▷

We will return later to a discussion analyzing the geometry of the realizations of a framework that give rise to special behaviors, such as Example 4.20 from Section 4.1. For now it suffices to define what we mean by "generic." In [**361**] a framework is *generic* if all of the coordinates of the joints are algebraically independent over $\mathbb{Q}$. Indeed, then we may treat these coordinates as indeterminates (or variables), which we denote by $\mathbf{x}$.

Like the proof of the LPG theorem, the combinatorial characterization of body-and-bar rigidity depends on the development of a rigidity matrix, R_G, and is stated in terms of infinitesimal motions.

Let $k = \binom{d+1}{2}$ be the dimension of the group of Euclidean motions. We embed $\mathbb{R}^d \subseteq \mathbb{P}^d$ via $(p_1, \ldots, p_d) \mapsto (1, p_1, \ldots, p_d)$ in the usual way and abuse notation letting $a_e = (1, a_e)$. In the linearization of the body-and-bar constraint equations given by White and Whiteley, the vector $\mathbf{p}_e$ of the Plücker coordinates of the line joining a_e and b_e replaces the displacement vectors in the rigidity matrix of the bar-and-joint frameworks.

Given a body-and-bar framework $G(\mathbf{p})$ with n vertices and m edges we form an $m \times kn$ rigidity matrix. If e has endpoints i and j with $i < j$, then the row corresponding to e has the form

$$(0 \cdots 0 \ \mathbf{p}_e \, 0 \cdots \ -\mathbf{p}_e \, 0 \cdots 0)$$

with $\mathbf{p}_e$ in the k columns associated to vertex i, $-\mathbf{p}_e$ in the k columns associated to vertex j, and zeroes in all other columns. The elements of the kernel of the rigidity matrix are precisely the infinitesimal motions of the framework.

White and Whiteley work in the more general setting of a k-frame.

Definition 4.59 (Definition 1.7 in [**361**]). A *k-frame* $G(\mathbf{p})$ consists of a multigraph $G = (V, E)$ and a vector $\mathbf{p}_e$ in $\mathbb{R}^k$ for each edge $e \in E$. We let $G(\mathbf{x})$ denote the generic k-frame associated to the graph G and interpret $\mathbf{x}$ as a vector of indeterminates.

Note that every body-and-bar framework is a k-frame in which the edge labels are Plücker coordinates of the lines along the bars of the framework. However, not all k-frames arise as body-and-bar frameworks.

We will focus our attention on the generically minimally rigid case where $m = kn - k$. As with bar-and-joint frameworks, we cannot eliminate the "trivial" motions of the framework viewed as a rigid body by inserting edges—these edges can only eliminate relative motions between joints. We eliminate trivial motions by tying down a body, say body one. In terms of the rigidity matrix we eliminate the infinitesimal motions associated to body one by adding k rows containing the $k \times k$ identity matrix I_k in the first k columns and zeroes elsewhere. We denote the

square matrix with tie-down by $R_{G,T}$:

$$R_{G,T} = \begin{bmatrix} \vdots & & & & & & \vdots \\ 0 & \cdots & \mathbf{x}_e & \cdots & -\mathbf{x}_e & \cdots & 0 \\ \vdots & & & & & & \vdots \\ I_k & 0 & \cdots & & & \cdots & 0 \end{bmatrix}.$$

DEFINITION 4.60. A k-frame $G(\mathbf{p})$ is *generically infinitesimally rigid* if the kernel of the matrix associated to $G(\mathbf{x})$ contains only the trivial infinitesimal motions. It is *minimally infinitesimally rigid* if the deletion of any edge of G produces a nontrivial infinitesimal motion.

THEOREM 4.61 ([**342**],[**361**]). *A generic body-and-bar framework associated to a multigraph $G(V, E)$ is minimally infinitesimally rigid in $\mathbb{R}^d$ if and only if*

(1) $|E| = k|V| - k$
(2) *If $V' \subseteq V$ is nonempty and $E' \subseteq E$ is the set of edges induced on V', then $|E'| \leq k|V'| - k$,*

where $k = \binom{d+1}{2}$.

DEFINITION 4.62 ([**361**]). Let G be a d-dimensional k-frame with $kn-k$ edges. We define $C_G(\mathbf{x}) = \det R_{G,T}$ to be the *pure condition* of G.

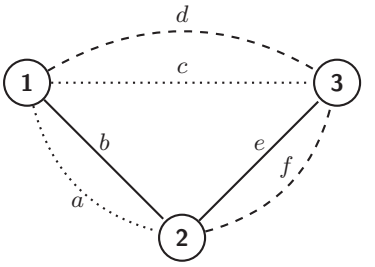
It is clear that if $G(\mathbf{p})$ is a k-frame with $kn - k$ edges, then it is minimally infinitesimally rigid if and only if the determinant of $R_{G,T}$ is nonzero. Moreover, the pure condition $C_G(\mathbf{x})$ associated to $G(\mathbf{x})$ is a nonzero polynomial if and only if G is minimally infinitesimally rigid. And, if G is generically minimally infinitesimally rigid i.e., $C_G(\mathbf{x})$ is nonzero, then $G(\mathbf{p})$ is minimally infinitesimally rigid for almost all $\mathbf{p}$ because the variety $\mathbf{V}(C_G)$ has measure zero. Theorem 4.61 gives a combinatorial characterization of this phenomenon.

THEOREM 4.63 ([**342**],[**361**]). *A generic body-and bar framework associated to a multigraph $G(V, E)$ is minimally infinitesimally rigid in $\mathbb{R}^d$ if and only if G can be decomposed as a union of k edge-disjoint spanning trees where $k = \binom{d+1}{2}$.*

Note that a tree on n vertices has $n - 1$ edges, so a graph that is the union of k edge-disjoint spanning trees has $kn - k$ edges. Hence, we see that if a graph has a decomposition into spanning trees as in Theorem 4.63, the first condition in Theorem 4.61 is satisfied.

EXAMPLE 4.64. The simplest nontrivial example of a generically minimally rigid body-and-bar framework with three bodies in the plane is given in Example 3.1 in [**361**]. A planar body-and-bar framework is a 3-frame, and we see that if we have two edges between each of the three bodies, the edges can be partitioned into three disjoint spanning trees (indicated below by dashed, dotted, and solid lines).

G =



We also give the rigidity matrix associated to G with vertex two tied down:

$$R_{G,T} = \left[\begin{array}{ccc|ccc|ccc} a_1 & a_2 & a_3 & -a_1 & -a_2 & -a_3 & 0 & 0 & 0 \\ b_1 & b_2 & b_3 & -b_1 & -b_2 & -b_3 & 0 & 0 & 0 \\ c_1 & c_2 & c_3 & 0 & 0 & 0 & -c_1 & -c_2 & -c_3 \\ d_1 & d_2 & d_3 & 0 & 0 & 0 & -d_1 & -d_2 & -d_3 \\ 0 & 0 & 0 & e_1 & e_2 & e_3 & -e_1 & -e_2 & -e_3 \\ 0 & 0 & 0 & f_1 & f_2 & f_3 & -f_1 & -f_2 & -f_3 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{array}\right].$$

We compute C_G by expanding the determinant via 3×3 minors which we denote as a bracket of three vectors. If we start with the expansion using the last three rows we see that $C_G = [abd][cef] - [abc][def]$. ◁▷

Note that the LPG theorem can be rephrased so that it is also in terms of trees.

THEOREM 4.65 ([**184**],[**262**]). *A generic bar-and-joint framework associated to $G(V, E)$ is minimally infinitesimally rigid in $\mathbb{R}^2$ if and only if adding any edge e to G (which may give a multigraph) results in a graph that can be decomposed into 2 edge-disjoint spanning trees.*

We will prove the more general statement for body-and-cad frameworks, and Theorem 4.63 will follow as a special case. The proof is essentially the same as the one given in [**361**] with minor modifications for the more general setting.

Body-and-Cad Frameworks. In the above discussion we defined body-and-bar frameworks which consist of a finite set of rigid bodies and distance constraints between points on pairs of bodies. We also saw that from the point of the combinatorial and algebraic analysis we could work more generally with k-frames. In this section we will introduce $[a, b]$-frames, a generalization of the k-frame that is a combinatorial model for other geometric constraints that can be imposed between a pair of rigid bodies.

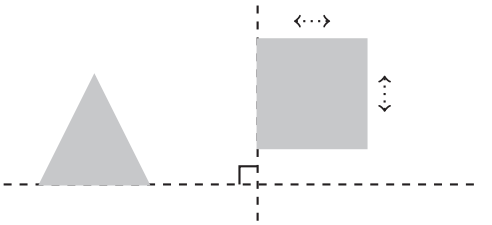
Given a pair of full-dimensional rigid bodies in $\mathbb{R}^d$ we specify geometric constraints between them by affixing affine linear spaces of dimension $0, \ldots, d-1$ to them and imposing coincidence, angle, or distance constraints between these linear spaces. For example, a "bar" in a body-and-bar framework is a *distance* constraint between two 0-dimensional linear spaces.

We give examples of other cad constraints in the plane below.

EXAMPLE 4.66. In the plane a constraint requiring two lines to coincide removes two degrees of freedom. Here we affix a line to each polygon and require the lines (dashed in the figure) to coincide. If we fix the triangle, then the square can translate horizontally but cannot rotate or translate in any other direction.

◁▷

EXAMPLE 4.67. An angle constraint between two lines in the plane restricts only one degree of freedom. The full 2-dimensional space of translations remains unrestricted by the requirement that the two dashed lines stay perpendicular.



◁▷

The authors Haller et al. [**185**] laid out a rigidity theory for cad (coincidence, angle, and distance) constraints with nine constraints in dimension two and 27 in dimension three. As in the body-and-bar setting, it is possible to linearize and work in the infinitesimal setting. We refer the reader interested in the derivation of a combinatorial model and rigidity matrix for body-and-cad frameworks to the original paper [**185**] or to [**330**]. The reader interested in the rigidity theory of other constraint systems may orient themselves by consulting [**235**] and [**219**].

We ask the same question as in the two previous cases, does a given set of constraints specify a structure that is generically infinitesimally rigid?

Before giving formal definitions it is important to note two key differences. First, a single geometric constraint may eliminate more than one degree of freedom as we saw in Example 4.66. Additionally, some geometric constraints only remove rotational degrees of freedom. For example, a system of constraints containing only line-line angle constraints cannot eliminate translational motions between the bodies. This is what happened in Example 4.67.

This information can be summarized in a *primitive cad graph*. This is a multigraph $H = (V, R \cup B)$ with one vertex for each rigid body. Each geometric constraint gives rise to one edge for each degree of freedom constrained by it, and the edges are partitioned into two sets, R (red) and B (black), with the red edges corresponding to constraints that only restrict rotational degrees of freedom. In dimension two there is a single angle degree of freedom, and in dimension three there are three angle degrees of freedom. As a consequence, some constraints are generically nonzero in all k-coordinates while others are forced to have zeroes in some coordinates, even generically.

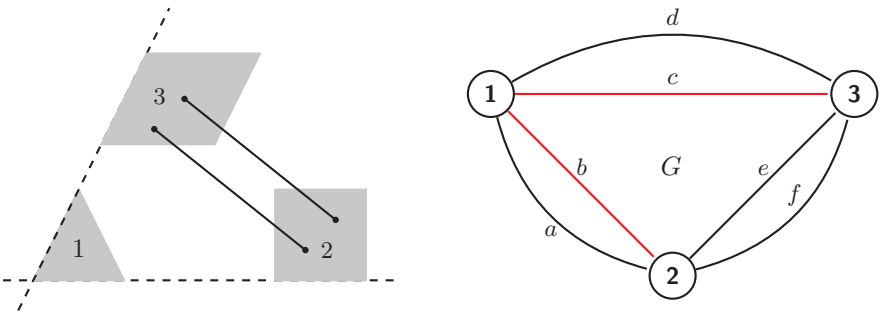
The definition of a body-and-cad framework, together with all of the possible constraints is given in [**185**]. For our purposes, the algebraic and combinatorial analysis of the linearized constraints, it suffices to define the purely combinatorial object, an $[a, b]$-frame, which allows us to keep track of angle constraints separately from "blind" constraints which are generically nonzero in all coordinates.

DEFINITION 4.68. An $[a,b]$-*frame* $G(\mathbf{x})$ is a bicolored multigraph $G=(V,R\cup B)$ together with a vector $(\mathbf{x}_e)$ of length $k=a+b$ for each edge, where the vectors labeling red edges have zeroes in the last b coordinates.

If $G(\mathbf{x})$ is an $[a,b]$-frame associated to a body-and-cad framework, then the underlying graph G is its primitive cad graph. Note that Definition 4.68 follows [**144**] which differs slightly from [**254**].

We will use the body-and-cad framework in Example 4.69, which is the case study in [**144**], to illustrate the theory throughout this section.

EXAMPLE 4.69. Let $G(\mathbf{x})$ be the $[1,2]$-frame associated to the body-and-cad framework consisting of a line-line coincidence constraint between bodies 1 and 2, a line-line coincidence constraint between bodies 1 and 3, and two point-point distance constraints between bodies 2 and 3 depicted below on the left, and the primitive cad graph G is on the right.



By Example 4.66, the line-line coincidence for bodies 1 and 2 restricts two degrees of freedom, one of which is rotational. This gives two edges in the graph G, with one colored red. The same holds for bodies 1 and 3. ◁▷

DEFINITION 4.70. $G=(V,R\cup B)$ is an $[a,b]$-*graph* if $\exists\, B'\subseteq B$ such that

(1) $(V,R\cup B')$ is the disjoint union of a spanning trees
(2) $(V,B\setminus B')$ is the disjoint union of b spanning trees

In Example 4.69 we see that G is actually a $[1,2]$-graph as the red edges form a spanning tree, and the remaining edges decompose into two other spanning trees (thus $B'=\emptyset$ in this case in Definition 4.70). Note that a body-and-bar framework is a special case of a body-and-cad framework, so it is possible to eliminate all of the relative motions of a body-and-cad framework with point-point distance constraints (bars). Thus, there are generically rigid body-and cad frameworks with no red edges. However, as we will see in the proof of Theorem 4.73, if we do have red edges then they will affect how we decompose G into spanning trees.

Associated to an $[a,b]$-frame $G(\mathbf{x})$ is a rigidity matrix R_G. As with the rigidity matrix of a k-frame, there are k columns per vertex. For each edge e from vertex i to vertex j with $i<j$, $\mathbf{x}_e$ appears in the k columns associated to vertex i and $-\mathbf{x}_e$ appears in the k columns associated to vertex j. As before, we can eliminate the trivial motions of the system by tying down the ith body. In the rigidity matrix this corresponds to appending k rows with a $k\times k$ identity matrix in the ith group of k columns and zeroes everywhere else.

If the $[a,b]$-frame $G(\mathbf{x})$ has $kn-k$ edges, then tying down a body results in squaring off the rigidity matrix and allows us to define a pure condition for $[a,b]$-frames.

DEFINITION 4.71. Given an $[a,b]$-frame $G(\mathbf{x})$ with $kn-k$ edges, its *pure condition* C_G is the determinant of $R_{G,T}$.

EXAMPLE 4.72. The tied-down rigidity matrix of the $[1,2]$-frame shown in Example 4.69 is

$$R_{G,T} = \left[\begin{array}{ccc|ccc|ccc} a_1 & a_2 & a_3 & -a_1 & -a_2 & -a_3 & 0 & 0 & 0 \\ b_1 & 0 & 0 & -b_1 & 0 & 0 & 0 & 0 & 0 \\ c_1 & 0 & 0 & 0 & 0 & 0 & -c_1 & 0 & 0 \\ d_1 & d_2 & d_3 & 0 & 0 & 0 & -d_1 & -d_2 & -d_3 \\ 0 & 0 & 0 & e_1 & e_2 & e_3 & -e_1 & -e_2 & -e_3 \\ 0 & 0 & 0 & f_1 & f_2 & f_3 & -f_1 & -f_2 & -f_3 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{array}\right]. \tag{4.26}$$

Red rows correspond to red edges in the graph, and being a $[1,2]$-frame gives extra zeros in the red rows. The pure condition is the determinant, which simplifies to

$$\begin{aligned} C_G = \det R_{G,T} &= \det \begin{bmatrix} a_1 & a_2 & a_3 & 0 & 0 & 0 \\ b_1 & 0 & 0 & 0 & 0 & 0 \\ c_1 & 0 & 0 & -c_1 & 0 & 0 \\ d_1 & d_2 & d_3 & -d_1 & -d_2 & -d_3 \\ 0 & 0 & 0 & -e_1 & -e_2 & -e_3 \\ 0 & 0 & 0 & -f_1 & -f_2 & -f_3 \end{bmatrix} \\ &= [abc][def] - [abd][cef] \\ &= -[abd][cef] \\ &= b_1(a_2d_3 - a_3d_2)c_1(e_2f_3 - e_3f_2). \end{aligned}$$

Interestingly, the two-term polynomial written in brackets is irreducible if the vectors are all generic 3-vectors. However, since b and c correspond to angle constraints, the first product of brackets vanishes. Moreover, the second product factors further because b and c correspond to angle constraints. ◁▷

THEOREM 4.73. *The pure condition of a generic $[a,b]$-frame $G(\mathbf{x})$ is nonzero if and only if it is an $[a,b]$-graph.*

We follow the proof of Theorem 2.18 in [**361**] closely, modifying it to accommodate the extra zeroes in the matrix in the rows corresponding to red edges. The direction showing C_G nonzero implies that G is an $[a,b]$-graph corrects what appears in [**254**]. The key to the proof is that the determinant of $R_{G,T}$ can be expanded via $n \times n$ minors or $k \times k$ minors in natural ways.

PROOF. First we will show that if $C_G(\mathbf{x})$ is nonzero, then the edges of G can be partitioned into k spanning trees as in Definition 4.70. Permute the columns of $R_{G,T}$ so that they are arranged in k groups of n, with all of the first coordinates of the vectors $\mathbf{x}_e$ in order, followed by all of the second coordinates, etc.

In the first a consecutive groupings of n columns we have a matrix whose first $kn-k$ rows are nonzero scalar multiples of the rows of the incidence matrix for the graph G with edges directed from low vertices to higher ones. In the last b groupings of n columns the rows are nonzero scalar multiples of the incidence matrix for the graph G minus all of its red edges (since red rows in these groupings are all zero).

For example, when we rearrange the columns of the matrix (4.26), the first six rows illustrate the structure just described:

(4.27)
$$\left[\begin{array}{ccc|ccc|ccc} a_1 & -a_1 & 0 & a_2 & -a_2 & 0 & a_3 & -a_3 & 0 \\ b_1 & -b_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ c_1 & 0 & -c_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ d_1 & 0 & -d_1 & d_2 & 0 & -d_2 & d_3 & 0 & -d_3 \\ 0 & e_1 & -e_1 & 0 & e_2 & -e_2 & 0 & e_3 & -e_3 \\ 0 & f_1 & -f_1 & 0 & f_2 & -f_2 & 0 & f_3 & -f_3 \\ \hline 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right].$$

If we expand this determinant via blocks of $n \times n$ minors, the determinant is a sum of products $\Pi_{i=1}^{k} \det M_i$ where each M_i is a subset of n rows in the ith block of columns chosen so that the union of the M_i partition the rows of $R_{G,T}$.

Since C_G is nonzero, at least one of these products must be nonzero. For such a product, each $\det M_i$ is nonzero. This means that the rows of M_i must not contain a cycle, and since we have chosen n rows, they must correspond to the $n-1$ edges of a tree and a tie-down row.

If $1 \leq i \leq a$, then the rows in M_i correspond to spanning trees that may contain red edges (together with some subset of black edges). The last b of the M_i cannot contain rows corresponding to red edges because these rows are entirely zero, as in (4.27). Therefore, we have found a decomposition of G into k spanning trees, a of which may contain red edges, and b that contain only black edges.

For the reverse direction we need to prove that if G is an $[a,b]$-graph, then the determinant of $R_{G,T}$ is nonzero. The plan is to produce a specialization of the generic edge labels at which C_G is nonzero, showing that it must be a nonzero polynomial.

Since G is an $[a,b]$-graph we have a trees that may contain red edges and b all black trees. Order them in such a way that the trees with red edges come first.

In each tree there is a unique path from any vertex v to the tie-down. Orient the edges of G so that all of these paths are directed towards the tie-down. Each of the remaining $n-1$ vertices has exactly k outgoing edges, one in each tree. The determinant of $R_{G,T}$ can be expanded as a sum of products of $k \times k$ minors in a natural way. As shown in the proof of Theorem 2.18 in [**361**], one of these, P, is the product of $k \times k$ minors consisting exactly of the rows corresponding to the k outgoing edges from each vertex, one from each tree

Now we give the edge labeling. Give each edge in tree i the same label, $(x_{i,1}, \ldots, x_{i,k})$, where the last b coordinates are zero for labels of red edges. If we evaluate P at this specialization of the edge labels, we see that we get a product of nonzero determinants. However, all of the other products of minors in the expansion of C_G must be zero. Indeed, to form any $k \times k$ minor we must choose k edges incident to each vertex. Any other choice requires that for some vertex we

choose two edges from the same tree, and since these edges have the same labeling in our specialization, that minor must be zero. □

As a consequence of Theorem 4.73, we get a combinatorial characterization of generic minimal infinitesimal rigidity for body-and-cad frameworks.

THEOREM 4.74 (Lee-St.John–S [**254**]). *A body-and-cad framework is generically minimally infinitesimally rigid if and only if it is a*

(1) $[1,2]$*-graph in 2D.*
(2) $[3,3]$*-graph in 3D* (*point-point coincidences are omitted*).

Interestingly, we must omit point-point coincidences from our constraint systems because the corresponding equations have symmetries that prevent us from incorporating them into the general theory above.

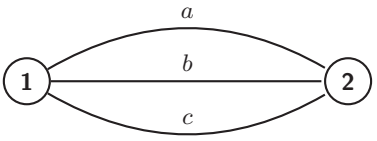
Nongeneric Frameworks. As we saw earlier, a k-frame $G(\mathbf{x})$ is generically minimally infinitesimally rigid if and only if $C_G(\mathbf{x})$ is a nonzero polynomial. If $\mathbf{p} \notin \mathbf{V}(C_G)$, then $G(\mathbf{p})$ is minimally infinitesimally rigid. Hence, almost all realizations of G have the generic behavior. Given this, it is natural to ask why we should study nongeneric frameworks.

The importance of nongeneric frameworks depends on the application. If the object of interest is truly a random realization of a given generic k-frame, then the combinatorial analysis suffices. However, frameworks designed to be useful are not random realizations of their combinatorics. They tend to have symmetries, repeated geometric patterns, and nongeneric parameters. Additionally, studying the nongeneric configurations of a framework allows us to deepen and complete our understanding of a framework as pure mathematicians.

We take the point of view of White and Whiteley in [**361**] and study nongeneric infinitesimal behavior by studying the pure condition C_G. This line of inquiry has two main branches. One is to use factors of C_G to understand important substructures of G. The other is to interpret the vanishing of a factor with an intuitive geometric description of the corresponding special positions.

To provide some intuition we return to Example 4.58.

EXAMPLE 4.75. The corresponding 3-frame $G(\mathbf{x}) = G(a,b,c)$ is depicted below.



The pure condition $C_G = [abc]$ vanishes whenever the points a, b, and c are collinear in the projective plane. But these points are the Plücker coordinates of the lines along the three bars, so in the dual projective space this corresponds to the condition that the three lines meet at a point. Here, we see an advantage of working in projective space. The two conditions that we identified initially, either the three lines are parallel or coincident, can be described by a single geometric condition in projective space because the case in which the three lines are parallel is equivalent to their common point of intersection lying on the line at infinity. ◁▷

We see that the vanishing of the pure condition in Example 4.75 has a nice geometric interpretation—the lines along the three bars are coincident in projective

space. This interpretation is easy to decipher because the pure condition is the vanishing of a single bracket, and we already knew a geometric interpretation for the vanishing of a determinant.

Interestingly, it is sometimes possible to give similar interpretations, in terms of linear spaces meeting unexpectedly, using the *Grassmann-Cayley* algebra. Given a vector space W, let $\bigvee W$ be the usual exterior algebra on V, though we denote the exterior product with a $\vee$ instead of the usual $\wedge$. If $w_1, \ldots, w_s \in W$, then $w_1 \vee \cdots \vee w_s$ is an *extensor*, and a typical element of $\bigvee W$ is a linear combination of extensors. Since an extensor is nonzero if and only if $w_1, \ldots, w_s$ form a basis for their span, it is natural to identify an extensor with a subspace of W and the product of two extensors that represent subspaces with trivial intersection represents the *join* of the subspaces.

The reason that we write the exterior product using $\vee$ instead of $\wedge$ is that the Grassmann-Cayley algebra is formed by introducing a second operation on $\bigvee W$ called the *meet*. The meet of two extensors u and v is a scalar times the extensor representing the intersection of the subspaces represented by u and v if this subspace has the expected dimension. If the intersection of two subspaces does not have the expected dimension, then their meet is zero.

To give a more precise definition we must first review the notion of a *shuffle*. A shuffle of $[s]$ into sets of size k and $s-k$ is a permutation $\sigma \in S_s$ satisfying the following: $\sigma(1) < \sigma(2) < \cdots < \sigma(k)$ and $\sigma(k+1) < \sigma(k+2) < \cdots < \sigma(s)$. We let $\text{sign}(\sigma)$ denote the sign of σ as a permutation.

Definition 4.76. Let W be a vector space of dimension r and $u = u_1 \vee \cdots \vee u_s$ and $v = v_1 \vee \cdots \vee v_t$ be extensors with $s + t \geq r$. Then we define

$$u \wedge v = \sum_{\sigma} \text{sign}(\sigma)[u_{\sigma(1)} \cdots u_{\sigma(r-t)} v_1 \cdots v_t] u_{\sigma(r-t+1)} \vee \cdots \vee u_{\sigma(s)},$$

where the sum is over all shuffles σ that split $[s]$ into sets of size $r-t$ and $s-(r-t)$.

For example, suppose that a, b, c, d, e, f are vectors in $\mathbb{R}^3$. Then $(c \vee d) \wedge (a \vee b) = [cab]d - [dab]c$. Moreover,

$$((c \vee d) \wedge (a \vee b)) \wedge (e \vee f) = [cab][def] - [dab][cef] = [abc][def] - [abd][cef].$$

The reader should note that $[abc][def] - [abd][cef]$ is the pure condition of the body-and-bar framework in Example 4.64. If a polynomial in brackets can be written as a single term in the Grassmann-Cayley algebra using only meets and joins (without summing), then we say that it has a *Grassmann-Cayley factorization*. Not all bracket polynomials have such factorizations, and even when such a factorization exists, it may be computationally intensive to produce it. (See [**332**] for details.) However, such a factorization does invite a geometric interpretation of the vanishing of the bracket polynomial.

In the case of Example 4.64 we consider a, b, c, d, e, f to be points in $\mathbb{P}^2$ and we get the condition that the lines ab, cd, and ef meet in a point. Since these vectors are the edge labels of a body-and-bar framework, they are the Plücker coordinates of the lines lying along the bars in the framework. To get an interpretation in the original projective space we dualize, and ab represents a point where the lines represented by a and b intersect, etc. Therefore, the condition is that the points ab, cd, and ef are collinear. The analysis of the framework in Example 4.64 given above appeared in [**361**], and illustrates one line of analysis of frameworks using the

pure condition in which special geometry configurations causing the pure condition to vanish can be identified via a Grassmann-Cayley factorization.

Another line of inquiry involves using irreducible factors of the pure condition to identify rigid substructures of a framework. In the case of body-and-bar frameworks White and Whiteley [**361**] lay out a beautiful correspondence between properties of the pure condition C_G and properties of the underlying graph G. For example,

(1) Each edge of G is in the support of exactly on irreducible factor of C_G.
(2) Each factor of C_G is the pure condition of a graph obtained by the deletion or contraction of edges of G along with the deletion of any isolated vertices.

For a general $[a, b]$-graph G, it is still the case that each edge is in the support of exactly one irreducible factor of C_G. However, irreducible factors of C_G may not be the pure condition of any $[a, a]$, $[b, b]$, or $[a + b, a + b]$-subgraph of G. The circuits of body-and-bar frameworks all have the form of a minimally rigid graph plus an extra edge, whereas the structure of circuits of $[a, b]$-graphs is mysterious, motivating the problem below.

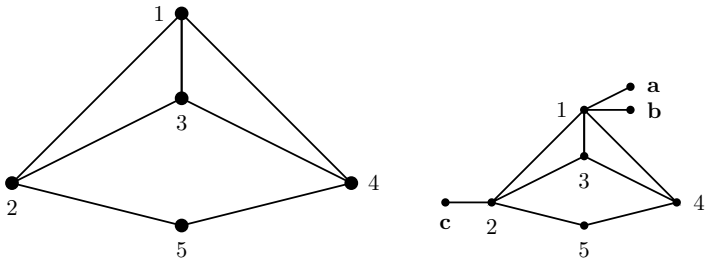
Problem: Describe and classify circuits of $[a, b]$-graphs.

Before their paper on body-and-bar frameworks, White and Whiteley introduced a notion of a pure condition for bar-and-joint frameworks in [**362**]. The pure condition for body-and-bar frameworks is presented first here because the correspondence between the polynomial and the combinatorics of the graph fits together so nicely—it really is a model for such a construction. Now we briefly sketch the theory for bar-and-joint frameworks, highlighting the differences between the two, before ending with a discussion of various ways of using the Cayley-Menger variety to study nongeneric realizations of frameworks.

Let $G(\mathbf{x})$ be a generically minimally infinitesimally rigid bar-and-joint framework. The pure condition, C_G, is defined as before by adding $\binom{d+1}{2}$ rows to the rigidity matrix to eliminate trivial motions of the framework as a rigid unit. However, in the case of bar-and-joint frameworks we work in the coordinates of the joints, and a single joint only has d coordinates. The interested reader can consult White and Whiteley [**362**], where they describe a general way of fixing joint coordinates to tie down the framework in a way that works well with the computation of the pure condition.

Instead of giving the general construction here we provide a detailed example.

Example 4.77. Consider the generic framework $G(\mathbf{x})$ associated to the graph on the left. We tie down the framework at vertices one and two, where the coordinates of the first joint will be completely determined by **a** and **b**.



The rigidity matrix associated to the tie-down is given below. The last three rows correspond to the tie-down. There are ten columns total as each column shown

represents a two-dimensional vector so is really two columns.

$$R_{G,T} = \left[\begin{array}{ccccc} \mathbf{x}_1-\mathbf{x}_2 & \mathbf{x}_2-\mathbf{x}_1 & 0 & 0 & 0 \\ \mathbf{x}_1-\mathbf{x}_3 & 0 & \mathbf{x}_3-\mathbf{x}_1 & 0 & 0 \\ \mathbf{x}_1-\mathbf{x}_4 & 0 & 0 & \mathbf{x}_4-\mathbf{x}_1 & 0 \\ 0 & \mathbf{x}_2-\mathbf{x}_3 & \mathbf{x}_3-\mathbf{x}_2 & 0 & 0 \\ 0 & \mathbf{x}_2-\mathbf{x}_5 & 0 & 0 & \mathbf{x}_5-\mathbf{x}_2 \\ 0 & 0 & \mathbf{x}_3-\mathbf{x}_4 & \mathbf{x}_4-\mathbf{x}_3 & 0 \\ 0 & 0 & 0 & \mathbf{x}_4-\mathbf{x}_5 & \mathbf{x}_5-\mathbf{x}_4 \\ \hline \mathbf{x}_1-\mathbf{a} & 0 & 0 & 0 & 0 \\ \mathbf{x}_1-\mathbf{b} & 0 & 0 & 0 & 0 \\ 0 & \mathbf{x}_2-\mathbf{c} & 0 & 0 & 0 \end{array}\right]$$

We expand $\det R_{G,T}$ via 2×2 minors, i.e., we multiply the determinants of two rows from the first block of two columns, two rows from the second two columns, etc. Note that if we view the joint coordinates as elements of projective space by sending $(x_{11}, x_{12}) \mapsto (1, x_{11}, x_{12})$, we can write a 2×2 minor as a 3×3 determinant in a nice way

$$\det \begin{bmatrix} \mathbf{x}_1-\mathbf{a} \\ \mathbf{x}_1-\mathbf{b} \end{bmatrix} = \det \begin{bmatrix} 1 & a_1 & a_2 \\ 1 & b_1 & b_2 \\ 1 & x_{11} & x_{12} \end{bmatrix}.$$

We will abuse notation and write the determinant above as the bracket $[\mathbf{a}, \mathbf{b}, \mathbf{x}]$. Hence, $\det R_{G,T} = [\mathbf{a}, \mathbf{b}, \mathbf{x}_1][\mathbf{x}_1, \mathbf{c}, \mathbf{x}_2][\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3][\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_4][\mathbf{x}_2, \mathbf{x}_4, \mathbf{x}_5]$. White and Whiteley show that with such a choice of tie down, the factors involving the tie-down variables are irrelevant, and we can define the pure condition of G to be $C_G = [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3][\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_4][\mathbf{x}_2, \mathbf{x}_4, \mathbf{x}_5]$.

Here, the vanishing of C_G has a natural geometric interpretation. Each factor is a 3×3 minor which vanishes when the corresponding three points are collinear. So, the framework has infinitesimal motions when any one of the three triangles, $\{1, 2, 3\}, \{1, 3, 4\}$, or $\{2, 4, 5\}$ degenerates. ◁▷

At a point $\mathbf{p}$ such that $C_G(\mathbf{p}) = 0$, there must be a linear dependence on the rows of the rigidity matrix $R_G(\mathbf{p})$. A nonzero dependence relation on the rows of R_G is called an *equilibrium stress*, and the *support* of the stress is the set of edges corresponding to rows with nonzero coefficient in the stress. The stress coefficients have a physical interpretation, so it is important to know the subgraph where a stress is supported.

Interestingly, the support of a stress associated to a given irreducible factor of C_G need not be related in obvious way to the structure of the factor as the example below shows.

Example 4.78. Continuing along from Example 4.77, note that in this case C_G factors as the product of three brackets, each of which vanishes when a triangle of three points degenerates to a line. Two of the triangles, on vertices $\{1, 2, 3\}$ and $\{1, 3, 4\}$ are subgraphs of G, and each of these supports a stress. However, the rigidity matrix for the induced subgraph on vertices $\{2, 4, 5\}$ is

$$\begin{bmatrix} 0 & \mathbf{x}_2-\mathbf{x}_5 & 0 & 0 & \mathbf{x}_5-\mathbf{x}_2 \\ 0 & 0 & 0 & \mathbf{x}_4-\mathbf{x}_5 & \mathbf{x}_5-\mathbf{x}_4 \end{bmatrix},$$

and we can see by inspection that these rows cannot support a stress alone. ◁▷

As Example 4.78 shows, irreducible factors of the pure condition can have the same algebraic structure while the combinatorics of the associated stresses is very different. To understand where the stresses associated to the vanishing of an irreducible factor of the pure condition are supported, we will return to the Cayley-Menger variety, which first made its appearance in Section 4.1. The advantage of working with the Cayley-Menger variety is that its defining equations are given in terms of the (squared) distances, which correspond to edges in the graph of the framework. Hence, the support of a polynomial naturally corresponds to a unique subgraph.

Recall that if we have n points in $\mathbb{R}^d$ then we can define $\varphi : (\mathbb{R}^d)^n \to \mathbb{R}^{\binom{n}{2}}$ by $\varphi_{ij}(\mathbf{x}) = (\mathbf{x}_i - \mathbf{x}_j) \cdot (\mathbf{x}_i - \mathbf{x}_j) = \ell_{ij}$, whose output is a vector of (squared) distances between pairs of points. The map φ was denoted $\ell_{K_n,d}$ in Section 4.1. Since we will use algebraic techniques, we regard φ as a map over $\mathbb{C}$, and we let $\varphi^* : \mathbb{C}[\ell_{ij}] \to \mathbb{C}[\mathbf{x}]$ be the map on rings corresponding to the length map. Given a graph G on n vertices, we can follow this map by the projection $\pi_E : \mathbb{C}^{\binom{n}{2}} \to \mathbb{C}^E$ given by the edges E of G. Function composition gives us the triangle of maps:

$$\begin{array}{ccc} (\mathbb{C}^d)^n & \xrightarrow{\quad\varphi\quad} & {}^aCM_{n,d,\mathbb{C}} \subseteq \mathbb{C}^{\binom{n}{2}} \\ & {\scriptstyle\varphi_G}\searrow \quad \swarrow{\scriptstyle\pi_E} & \\ & \mathbb{C}^E & \end{array}$$

Here, $\ker(\varphi^*)$ is the ideal of the Cayley-Menger variety ${}^aCM_{n,d,\mathbb{C}}$ and φ_G is the map $\ell_{G,d}$ from Section 4.1. A version of this diagram appeared as (4.14).

Each element of the ideal of the Cayley-Menger variety gives rise to a stress via the chain rule.

LEMMA 4.79. *If $f \in \ker(\varphi^*)$, then ∇f is a stress whose support is contained in the support of f.*

PROOF. By definition of $\ker(\varphi^*)$, $f(\varphi(\mathbf{x}))$ is identically zero. Hence,

$$d[f(\varphi(\mathbf{x}))] = \nabla f(\varphi(\mathbf{x})) \cdot d\varphi = 0.$$

Up to a constant, $d\varphi$ is the rigidity matrix for the complete graph on n vertices (see (4.7)). Hence the equation above shows that ∇f is a stress. It is clear that the support of ∇f is contained in the support of f. □

Via Lemma 4.79, we can see how to find a stress associated to each factor of the pure condition.

THEOREM 4.80 (Rosen–S–Theran–Vinzant). *Let $G = (V, E)$ be a minimally rigid graph in dimension d. If g is an irreducible factor of C_G, then*

(1) *$P = (\varphi^*)^{-1}(\langle g \rangle)$ is prime.*
(2) *$\exists\, f \neq 0 \in P \cap \mathbb{C}[\ell_{ij} \mid ij \in E]$.*
(3) *∇f is an equilibrium stress for frameworks in $\mathbf{V}(g)$.*

EXAMPLE 4.81. Continuing on with Example 4.77, using MACAULAY2 we can see that the defining equation of $\varphi_G(\mathbf{V}([\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3]))$ is given by

$$-\ell_{12}^2 + 2\ell_{12}\ell_{13} - \ell_{13}^2 + 2\ell_{12}\ell_{23} + 2\ell_{13}\ell_{23} - \ell_{23}^2,$$

which (up to sign) is the Cayley-Menger determinant from Example 4.26:

$$\det \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & \ell_{12} & \ell_{13} \\ 1 & \ell_{12} & 0 & \ell_{23} \\ 1 & \ell_{13} & \ell_{23} & 0 \end{bmatrix}.$$

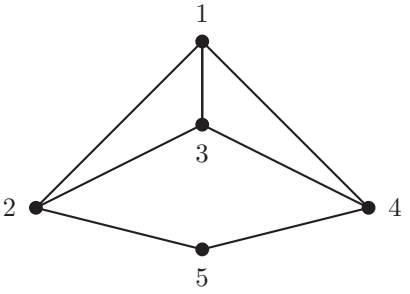
This is the Cayley condition that a triple of real numbers must satisfy if they are the pairwise distances among 3 collinear points. The defining equation of $\varphi_G(\mathbf{V}([\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_4]))$ is analogous. On the other hand, the defining equation of $\varphi_G(\mathbf{V}([\mathbf{x}_2, \mathbf{x}_4, \mathbf{x}_5]))$ is an irreducible polynomial of degree 6 supported on every edge of G. ◁▷

Motions of Frameworks. In the above discussion we saw how polynomials in the ideal of the Cayley-Menger variety give stresses on a framework. This is part of the infinitesimal analysis of rigidity. It is natural to ask if there is a way to use any of the algebra that we have set up to detect frameworks in positions where they actually move. This is the focus of ongoing work between the author and collaborators Zvi Rosen, Louis Theran, and Cynthia Vinzant.

We leave the reader with an example of a procedure for detecting motions:

Example 4.82 (Rosen–S–Theran–Vinzant). Let $G = (V, E)$ be the graph in the plane given below. The idea is that since G is generically minimally rigid, E is a maximal independent set in the algebraic matroid associated to $\ker(\varphi^*)$, which is the rigidity matroid $\mathcal{R}_2(5)$. If we add in a non-edge e, then $G + e$ must contain a circuit. We can view the associated circuit polynomial f_e as a polynomial in the variable e. For a general choice of the other edge lengths, we get finitely many possibilities for e. Letting I be the ideal of the coefficients of f_e, we see that the variety defined by I consists of edge lengths for which there are no conditions on e, so there must be infinitely many possibilities for it. Hence, a framework with these edge lengths has realizations that have motions.

- Add in edge $e = 35$.
- $G + e$ contains a circuit.
- Get circuit polynomial p of degree 4 in each variable.
- $I = \langle \text{coefficients of } p(e) \rangle$.
- I has 6 associated primes.



The two that are geometrically significant for motions are

$$\langle \ell_{13}, \ell_{14} - \ell_{34}, \ell_{12} - \ell_{23} \rangle, \langle \ell_{23} - \ell_{34}, \ell_{12} - \ell_{14}, -\ell_{45} + \ell_{25} \rangle.$$

What we see is that if vertices 1 and 3 coincide, $\ell_{14} = \ell_{34}$, and $\ell_{12} = \ell_{23}$ then a motion is possible. A realization of such a graph looks like a 4-bar mechanism.

The reader may check that the other ideal allows the graph to be realized with vertices 2 and 4 coinciding. This allows the doubled edge from vertices 2 and 4 to vertex 5 to rotate while the triangle is held fixed. ◁▷

Conclusion. Rigidity theory is an area that is ripe for modern algebraic methods. The infinitesimal combinatorial theory developed as a way to avoid working with the nonlinear (typically quadratic) polynomials that encode constraints. However, as we have seen, polynomials re-emerge in the combinatorial setting through the White-Whiteley pure condition. Moreover, working with polynomials in the length space and the Cayley-Menger variety presents new viewpoints and connections.

CHAPTER 5

Chemical Reaction Neworks

A network of chemical or biochemical reactions gives a system of differential equations by the Law of Mass Action. These systems have a lovely structure whose analysis involves graph theory and toric varieties.

5.1. The Classical Theory of Chemical Reactions

Our first goal is to make the transition from chemistry to mathematics. The starting point is the Law of Mass Action.

Law of Mass Action. Let's jump right in with a simple example.

EXAMPLE 5.1. In the atmosphere, the most common form of oxygen is the molecule O_2 consisting of two oxygen atoms. Nitric oxide NO and O_2 react to produce nitrogen dioxide NO_2. We write this reaction as

$$2\,\mathrm{NO} + \mathrm{O_2} \xrightarrow{\kappa} 2\,\mathrm{NO_2} \tag{5.1}$$

Let [NO], [O_2], [NO_2] denote concentrations, which are functions of time t. Naively, the concentration is the number of molecules per unit volume, though chemists use more natural units called *molarity*.

The *Law of Mass Action* describes the rate of change of concentration in a reaction. Applied to the above reaction, we get

$$\frac{d}{dt}[\mathrm{NO_2}] = 2\kappa[\mathrm{NO}]^2[\mathrm{O_2}], \tag{5.2}$$

We explain this as follows:

- The number of ways two NO and one O_2 can collide is roughly $[NO]^2[O_2]$.
- The *rate constant* κ corresponds to how many potential collisions actually occur and how fast the reaction is.
- The 2 on the right-hand side of the reaction (5.1) tells us that we preserve the number of atoms of N and O.
- This gives the 2 on the right-hand side of the differential equation (5.2).

There is a differential equation for each concentration, giving the system of ODEs:

$$\frac{d}{dt}\begin{bmatrix}[\mathrm{NO}]\\ [\mathrm{O_2}]\\ [\mathrm{NO_2}]\end{bmatrix} = \begin{bmatrix}-2\kappa[\mathrm{NO}]^2[\mathrm{O_2}]\\ -\kappa[\mathrm{NO}]^2[\mathrm{O_2}]\\ 2\kappa[\mathrm{NO}]^2[\mathrm{O_2}]\end{bmatrix} = \underbrace{\begin{bmatrix}2 & 0\\ 1 & 0\\ 0 & 2\end{bmatrix}}_{Y}\underbrace{\begin{bmatrix}-\kappa & 0\\ \kappa & 0\end{bmatrix}}_{A_\kappa}\underbrace{\begin{bmatrix}[\mathrm{NO}]^2[\mathrm{O_2}]\\ [\mathrm{N_2O}]^2\end{bmatrix}}_{\Phi}. \tag{5.3}$$

The product $YA_\kappa\Phi$ hints at the structure that underlies the reaction (5.1). ◁▷

A Small Lie. The chemistry of the reaction (5.1) is more complicated than first meets the eye. Our explanation of (5.2) required collisions between two NO

molecules and one O_2 molecule. But it is *extremely* rare for three molecules to collide simultaneously.

In reality, $2\,NO + O_2 \xrightarrow{\kappa} 2\,NO_2$ is the result of *elementary reactions* that involve at most two molecules. A proposed mechanism for $2\,NO + O_2 \xrightarrow{\kappa} 2\,NO_2$ uses two elementary reactions:

$$\begin{aligned} NO + NO &\underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} N_2O_2 \quad \text{(fast)} \\ N_2O_2 + O_2 &\xrightarrow{\kappa_3} 2\,NO_2 \quad \text{(slow)}. \end{aligned} \tag{5.4}$$

The first reaction reaches a steady state very quickly, so that

$$\kappa_1[NO]^2 = \kappa_2[N_2O_2]. \tag{5.5}$$

The second slower reaction then gives the differential equation

$$\frac{d}{dt}[NO_2] = 2\kappa_3[N_2O_2][O_2] = 2\kappa[NO]^2[O_2], \text{ where } \kappa = \kappa_3\,\frac{\kappa_1}{\kappa_2}.$$

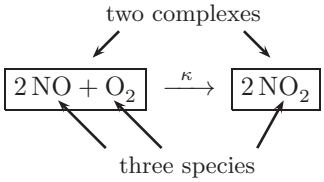
However, these reactions are so fast that chemists cannot be certain that this is what actually happens, which is why we say *proposed mechanism*. These and other subtleties are discussed in the 2012 article [**183**] by Gunawardena.

For us, the upshot is that when we apply the Law of Mass Action to a reaction like $2\,NO + O_2 \xrightarrow{\kappa} 2\,NO_2$, the resulting product $[NO]^2[O_2]$ is a convenient fiction that hides a more complicated reality.

Complexes, Species, and Directed Graphs. We now begin our study of the underlying structure of chemical reaction networks.

The molecules occurring in a reaction are called *species*, and combinations of species are called *complexes*. The reaction of Example 5.1 has two complexes and three species:

(5.6)



This reaction also gives some mathematical objects:

- For each species, its concentration is represented by a variable:

$$x_1 = [NO],\ x_2 = [O_2],\ x_3 = [NO_2].$$

- For each complex, we have a nonnegative linear combination of the species. This gives coefficient vectors:

$$y_1 = (2, 1, 0),\ y_2 = (0, 0, 2)$$

that tell us how each complex is built from the species. However, y_1 and y_2 are also the exponent vectors of the monomials:

$$[NO]^2[O_2] = x_1^2x_2 = x^{(2,1,0)} = x^{y_1},\ [NO_2]^2 = x_3^2 = x^{(0,0,2)} = x^{y_2}.$$

Then the system of differential equations (5.3) can be written more compactly as

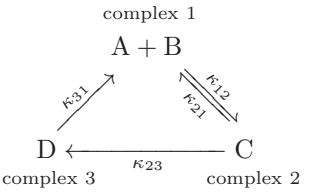
$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \kappa \underbrace{x_1^2 x_2}_{x^{y_1}} \left(\overbrace{\begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}}^{y_2} - \overbrace{\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}}^{y_1} \right). \tag{5.7}$$

Notice also that the two complexes in (5.6) form the vertices of the directed graph $\bullet \xrightarrow{\kappa} \bullet$ underlying the reaction.

We will soon see how all of this follows from the general form of the Law of Mass Action. But first, let's explore another example.

EXAMPLE 5.2. T-cell signal transduction gives the chemical reaction network

$$\begin{aligned} \mathrm{A} + \mathrm{B} &\underset{\kappa_{21}}{\overset{\kappa_{12}}{\rightleftharpoons}} \mathrm{C} \\ \mathrm{C} &\xrightarrow{\kappa_{23}} \mathrm{D} \\ \mathrm{D} &\xrightarrow{\kappa_{31}} \mathrm{A} + \mathrm{B} \end{aligned} \tag{5.8}$$



We have 4 species, 3 complexes, with 4 reactions since $\mathrm{A} + \mathrm{B} \underset{\kappa_{21}}{\overset{\kappa_{12}}{\rightleftharpoons}} \mathrm{C}$ is reversible. The directed graph has vertices given by the complexes and directed edges given by the reactions.

To reveal the structure of (5.8), we introduce variables and exponent vectors

$$\begin{aligned} &x_1 = [A],\ x_2 = [B],\ x_3 = [C],\ x_4 = [D] \\ &y_1 = (1,1,0,0),\ y_2 = (0,0,1,0),\ y_3 = (0,0,0,1). \end{aligned}$$

Then the Law of Mass Action gives the system of differential equations

$$\begin{aligned} \frac{dx_1}{dt} &= -\kappa_{12}x_1x_2 + \kappa_{21}x_3 + \kappa_{31}x_4 = -\kappa_{12}x^{y_1} + \kappa_{21}x^{y_2} + \kappa_{31}x^{y_3} \\ \frac{dx_2}{dt} &= -\kappa_{12}x_1x_2 + \kappa_{21}x_3 + \kappa_{31}x_4 = -\kappa_{12}x^{y_1} + \kappa_{21}x^{y_2} + \kappa_{31}x^{y_3} \\ \frac{dx_3}{dt} &= \kappa_{12}x_1x_2 - \kappa_{21}x_3 - \kappa_{23}x_3 = \kappa_{12}x^{y_1} - (\kappa_{21}+\kappa_{23})x^{y_2} \\ \frac{dx_4}{dt} &= \kappa_{23}x_3 - \kappa_{31}x_4 = \kappa_{23}x^{y_2} - \kappa_{31}x^{y_3}, \end{aligned} \tag{5.9}$$

which can be written

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \kappa_{12}x^{y_1} \underbrace{\begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix}}_{y_2-y_1} + \kappa_{21}x^{y_2} \underbrace{\begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \end{bmatrix}}_{y_1-y_2} + \kappa_{23}x^{y_2} \underbrace{\begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \end{bmatrix}}_{y_3-y_2} + \kappa_{31}x^{y_3} \underbrace{\begin{bmatrix} 1 \\ 1 \\ 0 \\ -1 \end{bmatrix}}_{y_1-y_3}. \tag{5.10}$$

What seemed like random + and − signs in (5.9) now make perfect sense in (5.10) once we realize that the column vectors on the right are differences $y_i - y_j$ for suitable i and j. ◁▷

Chemical Reaction Networks. A general chemical reaction network consists of the following data:

- n species and m complexes.
- a directed graph $G = (V, E)$ with positive weights $\kappa = (\kappa_{i\to j})$:
  - The vertices are the m complexes, i.e., $V = [m]$.
  - The directed edges are the reactions.
  - The weight of the directed edge $i\to j$ is the reaction rate $\kappa_{i\to j}$.
- An $n \times m$ matrix Y whose columns $y_1, \dots, y_m$ express the complexes as nonnegative integer linear combinations of the species.

Since a chemical reaction network is determined by its reactions, we assume that G has no isolated vertices. For simplicity, we often write κ_{ij} instead of $\kappa_{i\to j}$.

Given a chemical reaction network (G, κ, Y), let x_i be the concentration of the ith species. Each x_i is a function of time, and the Law of Mass Action tells us that the vector $x = (x_1, \dots, x_n)$ satisfies the system of ordinary differential equations

$$\frac{dx}{dt} = \sum_{i\to j\in E} \kappa_{i\to j}\, x^{y_i}\,(y_j - y_i), \quad x = (x_1, \dots, x_n)^t. \tag{5.11}$$

For example, the reaction (5.1) gives the system (5.7), and in Example 5.2, the reaction network (5.8) gives (5.10). Both of these are instances of the Law of Mass Action. In practice, the system (5.11) is often referred to as *mass-action kinetics*.

The Graph Laplacian. Equation (5.11) is not the only way to write the Law of Mass Action. Adapting (5.3) to our current notation, the systems of ODEs in Example 5.1 becomes

$$\frac{dx}{dt} = \underbrace{\begin{bmatrix} 2 & 0 \\ 1 & 0 \\ 0 & 2 \end{bmatrix}}_{Y} \underbrace{\begin{bmatrix} -\kappa & 0 \\ \kappa & 0 \end{bmatrix}}_{A_\kappa} \underbrace{\begin{bmatrix} x^{y_1} \\ x^{y_2} \end{bmatrix}}_{\Phi(x)}.$$

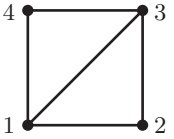
We now know what Y means, and $\Phi(x)$ is the column vector of monomials x^{y_i}. But what is the matrix A_κ in the middle? This is where the *graph Laplacian* enters the picture.

For a simple graph $G = ([m], E)$, the graph Laplacian is the $m \times m$ symmetric matrix $L = (L_{ij})$, where

$$L_{ij} = \begin{cases} \deg(i) & i = j \\ -1 & ij \in E \\ 0 & \text{otherwise.} \end{cases}$$

This matrix is symmetric with rows and columns that sum to zero.

Example 5.3. Here is a graph from Chapter 4 and its graph Laplacian



$$L = \begin{bmatrix} 3 & -1 & -1 & -1 \\ -1 & 2 & -1 & 0 \\ -1 & -1 & 3 & -1 \\ -1 & 0 & -1 & 2 \end{bmatrix}$$

◁▷

This idea adapts easily to weighted graphs, where the off-diagonal entries are now $-\kappa_{ij}$ and the diagonal entries are determined by the condition that the row and column sums are zero.

In our situation, we have a weighted directed graph (G, κ), which introduces a measure of asymmetry. Here, the *weighted directed graph Laplacian* is the $m \times m$ matrix L_κ defined by

$$(L_\kappa)_{ij} = \begin{cases} \displaystyle\sum_{i \to k \in E} \kappa_{i\to k} & i = j \\ -\kappa_{j\to i} & i \neq j,\ j{\to}i \in E \\ 0 & i \neq j,\ j{\to}i \notin E, \end{cases} \tag{5.12}$$

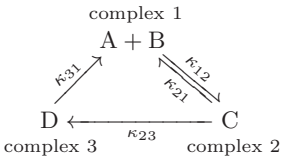
where the diagonal entries ensure that the column sums are zero. Here, we follow [**123**], where L_κ is denoted $L_{2,\kappa}$.

To get the matrix A_κ we want, we take the negative of L_κ, i.e., $A_\kappa = -L_\kappa$. Note that the column sums of A_κ are zero. With this definition of A_κ, the Law of Mass Action (5.11) can be written more compactly as

$$\frac{dx}{dt} = Y A_\kappa \Phi(x), \tag{5.13}$$

where Y is the $n \times m$ matrix whose columns $y_1, \ldots, y_m$ express the m complexes in terms of the n species, A_κ is as above, and $\Phi(x)$ is the column vector of monomials $x^{y_1}, \ldots, x^{y_m}$ in the species concentrations $x_1, \ldots, x_n$.

Example 5.4. Consider the chemical reaction network of Example 5.2:



We can write the system (5.9) in the form

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{Y} \underbrace{\begin{bmatrix} -\kappa_{12} & \kappa_{21} & \kappa_{31} \\ \kappa_{12} & -\kappa_{21} - \kappa_{23} & 0 \\ 0 & \kappa_{23} & -\kappa_{31} \end{bmatrix}}_{A_\kappa} \underbrace{\begin{bmatrix} x_1 x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{\Phi(x)}.$$

◁▷

Balancing Equations. There is another way to see how A_κ arises naturally from $G = ([m], E, \kappa)$. Suppose we have a quantity $\mathcal{C}_i$ at each vertex i. If we allow the $\mathcal{C}_i$ to flow along the directed edges with rates given by $\kappa_{i\to j}$, then

$$\begin{aligned} \text{inflow at vertex } i &= \sum_{j \to i \in E} \kappa_{j\to i}\, \mathcal{C}_j \\ \text{outflow at vertex } i &= \sum_{i \to j \in E} \kappa_{i\to j}\, \mathcal{C}_i = \Big(\sum_{i \to j \in E} \kappa_{i\to j} \Big) \mathcal{C}_i. \end{aligned}$$

The system is *balanced* when inflow = outflow at every vertex i, i.e., when

$$\sum_{j \to i \in E} \kappa_{j\to i}\, \mathcal{C}_j - \Big(\sum_{i \to k \in E} \kappa_{i\to j} \Big) \mathcal{C}_i = 0$$

for all i. In terms of the matrix A_κ, this system can be written

$$A_\kappa \begin{bmatrix} \mathcal{C}_1 \\ \vdots \\ \mathcal{C}_m \end{bmatrix} = 0. \tag{5.14}$$

These *balancing equations* will play an important role in what follows.

Stoichiometry. The Greek words στοιχείο (stoicheío) and μέτρο (métro) translate to *element* and *measure* respectively. Chemical reactions neither create nor destroy atoms; stoichiometry is the quantitative consequence of this basic principle.

EXAMPLE 5.5. The reaction $2\,\mathrm{NO} + \mathrm{O}_2 \xrightarrow{\kappa} 2\,\mathrm{NO}_2$ has conserved quantities:

(Nitrogen atoms) $$[\mathrm{NO}] + [\mathrm{NO}_2] = [\mathrm{NO}]_0 + [\mathrm{NO}_2]_0$$

(Oxygen atoms) $$[\mathrm{NO}] + 2\,[\mathrm{O}_2] + 2\,[\mathrm{NO}_2] = [\mathrm{NO}]_0 + 2\,[\mathrm{O}_2]_0 + 2\,[\mathrm{NO}_2]_0$$

The subscript 0 denotes initial concentrations. After some algebra, one obtains

$$\begin{bmatrix} [\mathrm{NO}] \\ [\mathrm{O}_2] \\ [\mathrm{NO}_2] \end{bmatrix} = \begin{bmatrix} [\mathrm{NO}]_0 \\ [\mathrm{O}_2]_0 \\ [\mathrm{NO}_2]_0 \end{bmatrix} + \lambda(t) \begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix}, \ \lambda(t) = -\tfrac{1}{2}([\mathrm{NO}] - [\mathrm{NO}]_0).$$

For this reaction, stoichiometry implies that the reaction evolves along a line. ◁▷

Looking deeper into this example reveals the structure more clearly. When we write the Law of Mass Action in the form (5.11) for this reaction, we get

$$\frac{dx}{dt} = \kappa\, x^{y_1}(y_2 - y_1), \quad y_2 - y_1 = \begin{bmatrix} 0-2 \\ 0-1 \\ 2-0 \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix}.$$

The tangent vector $\dfrac{dx}{dt}$ always lies in the span of $y_2 - y_1$, so that

$$x(t) \in x(0) + \mathrm{Span}(y_2 - y_1). \tag{5.15}$$

A key observation is that the stoichiometry is determined by $y_2 - y_1$. More precisely:

- By (5.15), any vector v satisfying $v \cdot (y_2 - y_1) = 0$ has the property that $v \cdot x(t) = v \cdot x(0)$ for all t. Hence $v \cdot x$ is a conserved quantity.
- One computes that $\{(1,0,1),(1,2,2)\}$ is a basis of $(y_2 - y_1)^\perp$. This gives another way to see that $[\mathrm{NO}]+[\mathrm{NO}_2]$ and $[\mathrm{NO}]+2\,[\mathrm{O}_2]+2\,[\mathrm{NO}_2]$ are conserved.

Thus everything in Example 5.5 is a consequence of the vector $y_2 - y_1$.

The Stoichiometric Subspace. For a general mass-action kinetics system

$$\frac{dx}{dt} = \sum_{i \to j \in E} \kappa_{i\to j}\, x^{y_i}\,(y_j - y_i), \quad x = (x_1, \ldots, x_n)^t.$$

as in (5.11), the *stoichiometric subspace* is

$$S = \mathrm{Span}(y_j - y_i \mid i \to j \in E).$$

If $x = x(t)$ is a solution of the system, then arguing as above shows that

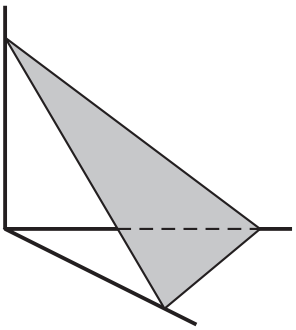
$$x(t) \in x(0) + S.$$

It follows that if we have a vector $v \in S^\perp$, then $v \cdot x$ is a conserved quantity and equals the constant $v \cdot x(0)$.

However, in chemical reactions, concentrations are *nonnegative*. Thus $x(t) \in \mathbb{R}^n_{\geq 0}$, which implies that

$$x(t) \in (x(0) + S) \cap \mathbb{R}^n_{\geq 0} \text{ for } t \geq 0. \tag{5.16}$$

We call $(x(0) + S) \cap \mathbb{R}^n_{\geq 0}$ a *stoichiometric compatibility class*. Here is a picture:



In general, stoichiometric compatibility classes are polyhedra. The above example is bounded, i.e., a polytope. In Section 5.2 we will see an example of an unbounded stoichiometric compatibility class.

Above, we justified (5.16) by saying that in the real world, concentrations cannot be negative. This can also be proved mathematically, i.e.,

$$x(0) \in \mathbb{R}^n_{\geq 0} \Longrightarrow x(t) \in (x(0) + S) \cap \mathbb{R}^n_{\geq 0} \text{ for } t \geq 0.$$

A proof can be found in [**129**].

Steady State Solutions. An important idea in chemistry is that most chemical reactions achieve an equilibrium as time goes on. Let's write the corresponding system in the form (5.13)

$$\frac{dx}{dt} = Y A_\kappa \Phi(x).$$

A *steady state* or *equilibrium solution* is a solution where the concentration vector $x = (x_1, \ldots, x_n)^t$ is constant. Thus x is a steady state solution if and only if

$$Y A_\kappa \Phi(x) = 0.$$

This system of polynomial equations shows that the steady state solutions lie in the *steady state variety*

$$V^{ss} = \mathbf{V}(Y A_\kappa \Phi(x)) \subseteq \mathbb{R}^n. \tag{5.17}$$

The physically meaningful steady state solutions correspond to

$$V^{ss}_{\geq 0} = \mathbf{V}(Y A_\kappa \Phi(x)) \cap \mathbb{R}^n_{\geq 0}. \tag{5.18}$$

Observe that V^{ss} is a real algebraic variety and $V^{ss}_{\geq 0}$ is a semialgebraic variety. Here are two simple examples.

EXAMPLE 5.6. For $2\,\mathrm{NO} + \mathrm{O}_2 \xrightarrow{\kappa} 2\,\mathrm{NO}_2$, the steady states of the system (5.3) satisfy $[\mathrm{NO}]^2[\mathrm{O}_2] = 0$, so

$$[\mathrm{NO}] = 0 \text{ or } [\mathrm{O}_2] = 0.$$

In other words, we run out of one of the reactants. This is *extinction*. ◁▷

EXAMPLE 5.7. In (5.4), we decomposed Example 5.6 into elementary reactions, one of which is $\mathrm{NO} + \mathrm{NO} \underset{\kappa_{21}}{\overset{\kappa_{12}}{\rightleftharpoons}} \mathrm{N_2O_2}$. The resulting system is

$$\begin{aligned} \tfrac{d}{dt}[\mathrm{NO}] &= -2\kappa_{12}[\mathrm{NO}]^2 + 2\kappa_{21}[\mathrm{N_2O_2}] \\ \tfrac{d}{dt}[\mathrm{N_2O_2}] &= \quad \kappa_{12}[\mathrm{NO}]^2 - \ \kappa_{21}[\mathrm{N_2O_2}]. \end{aligned}$$

The steady states satisfy the equation

$$\kappa_{12}[\mathrm{NO}]^2 = \kappa_{21}[\mathrm{N_2O_2}]$$

that appeared earlier in (5.5). ◁▷

Three Examples. So far, we have studied networks of chemical reactions. But there is a wide range of phenomena having nothing to do with chemistry that can modeled using "chemical reaction networks". Here we present three examples, one of which we've already seen.

EXAMPLE 5.8. Recall the HIV model from [**179**] considered in Example 2.16:

Description	Interaction	Parameter value
Generation of new T cells	$\emptyset \xrightarrow{s_1} T$	10
Generation of new macrophages	$\emptyset \xrightarrow{s_2} M$	1.5×10^{-1}
Proliferation of T cells by presence of pathogen	$T + V \xrightarrow{k_1} (T+V) + T$	2×10^{-3}
Infection of T cells by HIV	$T + V \xrightarrow{k_2} T_i$	3×10^{-3}
Proliferation of M by presence of pathogen	$M + V \xrightarrow{k_3} (M+V) + M$	7.45×10^{-4}
Infection of M by HIV	$M + V \xrightarrow{k_4} M_i$	5.22×10^{-4}
Proliferation of HIV within T cell	$T_i \xrightarrow{k_5} V + T_i$	5.37×10^{-1}
Proliferation of HIV within macrophage	$M_i \xrightarrow{k_6} V + M_i$	2.85×10^{-1}
Natural death of T cells	$T \xrightarrow{\delta_1} \emptyset$	0.01
Natural death of infected T cells	$T_i \xrightarrow{\delta_2} \emptyset$	0.44
Natural death of macrophages	$M \xrightarrow{\delta_3} \emptyset$	6.6×10^{-3}
Natural death of infected macrophages	$M_i \xrightarrow{\delta_4} \emptyset$	6.6×10^{-3}
Natural death of HIV	$V \xrightarrow{\delta_5} \emptyset$	?

The interactions between cells shown here clearly form a chemical reaction network. One difference is that cells can be created and can die. For example, the "reactions" $\emptyset \xrightarrow{s_1} T$ and $T \xrightarrow{\delta_1} \emptyset$ represent the creation and death of T cells.

We think of these interactions as a *biochemical network*, where the species are cells rather than atoms. Applying the Law of Mass Action, this network gives

$$\begin{aligned} \tfrac{d}{dt}T &= s_1 + k_1TV - k_2TV - \delta_1 T \\ \tfrac{d}{dt}T_i &= k_2TV - \delta_2 T_i \\ \tfrac{d}{dt}M &= s_2 + k_3MV - k_4MV - \delta_3 M \\ \tfrac{d}{dt}M_i &= k_4MV - \delta_4 M_i \\ \tfrac{d}{dt}V &= k_5T_i + k_6M_i - \delta_5 V, \end{aligned} \tag{5.19}$$

exactly as in (2.23). It is nice to see how this biological example from [**179**] fits into the framework of chemical reaction networks. ◁▷

EXAMPLE 5.9. In epidemiology, the SIR model is given by

$$\begin{aligned} \text{S} &= \text{number of susceptible individuals} & \tfrac{d\text{S}}{dt} &= -\alpha \text{SI} \\ \text{I} &= \text{number of infected individuals} & \tfrac{d\text{I}}{dt} &= \alpha \text{SI} - \beta \text{I} \\ \text{R} &= \text{number of recovered individuals} & \tfrac{d\text{R}}{dt} &= \beta \text{I} \end{aligned}$$

Many courses in differential equations present this as a *compartment model*:

$$\boxed{\text{S}} \longrightarrow \boxed{\text{I}} \longrightarrow \boxed{\text{R}}.$$

Sometimes there are labels on the arrows:

$$\underbrace{\boxed{\text{S}} \xrightarrow{\alpha \text{SI}} \boxed{\text{I}} \xrightarrow{\beta \text{I}} \boxed{\text{R}}}_{\text{labels are rates of change}} \quad \text{or} \quad \underbrace{\boxed{\text{S}} \xrightarrow{\alpha \text{I}} \boxed{\text{I}} \xrightarrow{\beta} \boxed{\text{R}}}_{\text{labels are relative rates of change}}.$$

We can also think of this as a reaction network:

$$\text{S} + \text{I} \xrightarrow{\alpha} \text{I} \xrightarrow{\beta} \text{R}.$$

The Law of Mass Action gives the above differential equations.

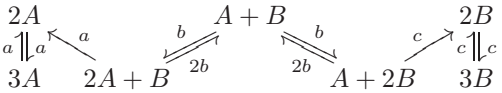
The difference between these two presentations of the SIR model is easy to understand: the reaction network is *complex based*, while the compartment model is *species based*. ◁▷

EXAMPLE 5.10. In population genetics, chromosomes have locations where different alleles can occur. Examples of alleles include blood type or the color of a flower. Assume that we have two alleles A, B with frequencies p, q. In this situation, the genotypes $AA, AB = BA, BB$ have what are called *Malthusian fitnesses* a, b, c. Roughly speaking, these are the logs of the average numbers of offspring.

Going from one generation to the next gives a discrete model. Passing to a continuous model, the resulting differential equations are

$$\begin{aligned} \tfrac{d}{dt}p &= p(\overbrace{ap + bq}^{\substack{\text{marginal} \\ \text{fitness of } A}}) - p(\overbrace{ap^2 + 2bpq + cq^2}^{\substack{\text{mean fitness of} \\ \text{the population}}}) \\ \tfrac{d}{dt}q &= q(bp + cq) - q(ap^2 + 2bpq + cq^2). \end{aligned}$$

In 1979, Hárs and Tóth [**189**] characterized which systems of ODEs come from chemical reaction networks. The above differential equations satisfy their criteria, and using the method described in [**189**], it is straightforward to show that these ODEs come from the network

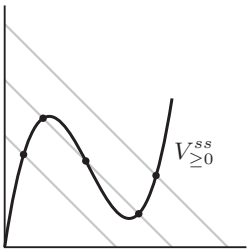


via the Law of Mass Action. ◁▷

Questions About Steady States.. Given a system $\frac{dx}{dt} = YA_\kappa \Phi(x)$, there are several questions to ask about a nonnegative steady state solution $x_* \in V_{\geq 0}^{ss}$.

The stoichiometric compatibility class of x_* is $(x_* + S) \cap \mathbb{R}^n_{\geq 0}$, where $S \subseteq \mathbb{R}^n$ is the stoichiometric subspace. The first question is whether $(x_* + S) \cap \mathbb{R}^n_{\geq 0}$ contains

other steady states. In other words, how big is $V^{ss}_{\geq 0} \cap (x_* + S)$? The answer can depend on which stoichiometric compatibility class we are in, as illustrated here:



A network exhibits *multistationarity* when some stoichiometric compatibility classes contain more than one steady state. *Monostationarity* occurs where there is at most one steady state in each stoichiometric compatibility class.

Since solutions of $\frac{dx}{dt} = YA_\kappa\Phi(x)$ always stay within a given compatibility class, there are two further questions to ask about a steady solution $x_* \in V^{ss}_{\geq 0}$:

- Is x_* *locally attracting*, i.e., if $x(0) \in (x_* + S) \cap \mathbb{R}^n_{\geq 0}$ is close to x_*, does $x(t) \to x_*$ as $t \to \infty$?
- If $V^{ss}_{\geq 0} \cap (x_* + S) = \{x_*\}$, is x_* *globally attracting*, i.e., if $x(0) \in (x_* + S) \cap \mathbb{R}^n_{\geq 0}$, does $x(t) \to x_*$ as $t \to \infty$?

We will see eventually that some of these questions have very satisfying answers, while others remain open.

EXAMPLE 5.11. In the HIV model presented in Example 5.8, the steady states of (5.19) are solutions of the polynomial system

$$\begin{aligned}
0 &= s_1 + k_1TV - k_2TV - \delta_1 T \\
0 &= k_2TV - \delta_2 T_i \\
0 &= s_2 + k_3MV - k_4MV - \delta_3 M \\
0 &= k_4MV - \delta_4 M_i \\
0 &= k_5T_i + k_6M_i - \delta_5 V.
\end{aligned}$$

In Example 5.8, we assumed that all rate constants are known except for δ_5. The five types of cells T, T_i, M, M_i, V (the species) and the unknown rate constant δ_5 give the steady state variety $V^{ss} \subseteq \mathbb{R}^5 \times \mathbb{R} = \mathbb{R}^6$.

We saw in Example 2.16 that Gross and Bates [**179**] use MACAULAY2 to write the variety V^{ss} as the union of two irreducible components: the main component $V_m = \mathbf{V}(f_1, f_2, f_3, f_4, f_5)$ with

$$\begin{aligned}
f_1 &= 5742\,M - 2453\,M_i - 130500 \\
f_2 &= 259908\,T_i - 46607\,M_i + 4840000\,\delta_5 - 20200500 \\
f_3 &= 17721\,T + 46607\,M_i - 4840000\,\delta_5 + 2479500 \\
f_4 &= 484000\,V\delta_5 - 184547\,M_i + 4840000\,\delta_5 - 20200500 \\
f_5 &= 2453\,M_iV - 72600\,M_i + 130500\,V,
\end{aligned}$$

and the extinction component

$$V_e = \mathbf{V}(V, M_i, 11\,M - 250, T_i, T - 1000).$$

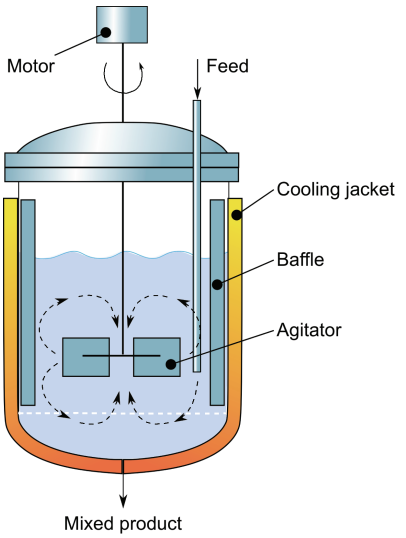
This is a nice example of how ideas from computational algebraic geometry can help explain a biological system. ◁▷

The Creators of Chemical Reaction Network Theory. Before discussing some of the classic results of CRNT, let's say a few words about the people behind this theory:

- Fritz Horn (1927–1978)
- Roy Jackson (1931–)
- Martin Feinberg (1942–)

Among their many papers on CRNT, most written between 1972 and 1995, are [**147**, **211**–**213**]. Horn, Jackson and Feinberg all came from chemical engineering, where a basic object is a *stirred tank*:

(5.20)



The cooling jacket is important, since reaction rates can vary with temperature but we want the reaction rates κ to be constant. The arrows labelled *Feed* and *Mixed product* allow reactants to be added and products to be removed. This corresponds to reactions $\emptyset \longrightarrow A$, $B \longrightarrow \emptyset$ in the network.

There are several excellent expositions of CRNT, including the 1979 lectures of Feinberg [**148**] (freely available on the internet) and the 2003 lectures [**182**] and 2012 survey [**183**] by Gunawardena. The latter has a nice discussion of the Matrix Tree Theorem, which will play a significant role in Section 5.2.

Two Classic Results of CRNT. The remainder of this section will be devoted to a discussion of two cornerstones of CRNT, the *Deficiency Zero Theorem* and the *Balanced Dynamics Theorem*. Here is a stripped-down version of the former.

THEOREM 5.12 (Deficiency Zero). *A chemical reaction network is* ***complex balanced*** *for every set of positive reaction rates κ if and only if it has* ***deficiency zero*** *and is* ***weakly reversible****.*

Once we explain "complex balanced" and "weakly reversible," we will be able to state the Balanced Dynamics Theorem. After defining "deficiency", we will be ready for the full version of the Deficiency Zero Theorem.

Complex Balanced Networks. A chemical reaction network (G, κ, Y) leads to two sorts of objects:

- Steady states x_* of the ODE system satisfy the equation $YA_\kappa\Phi(x_*) = 0$.
- Balancing equations of weighted directed graph (G, κ) have coefficient matrix A_κ by (5.14). These are the *complex balancing equations* since the vertices of G correspond to complexes.

A steady state x_* is *complex balanced* when $\Phi(x_*)$ satisfies the complex balancing equations, i.e., $A_\kappa\Phi(x_*) = 0$.

The following example shows that a complex balanced steady state need not use the whole network.

EXAMPLE 5.13. Consider the reaction network with complexes C_1, C_2, C_3, C_4:

$$\underbrace{C_1 \underset{\kappa_{21}}{\overset{\kappa_{12}}{\rightleftarrows}} C_2}_{\text{strong linkage class}} \xrightarrow{\kappa_{23}} \underbrace{C_3 \underset{\kappa_{43}}{\overset{\kappa_{34}}{\rightleftarrows}} C_4}_{\substack{\text{terminal} \\ \text{strong linkage class}}} \tag{5.21}$$

For the moment, ignore the labels in the diagram. The first two complex balancing equations of the weighted directed graph (5.21) are

$$\overset{\text{inflow } C_1}{\kappa_{21}\mathcal{C}_2} = \overset{\text{outflow } C_1}{\kappa_{12}\mathcal{C}_1} \quad \text{and} \quad \overset{\text{inflow } C_2}{\kappa_{12}\mathcal{C}_1} = \overset{\text{outflow } C_2}{\kappa_{21}\mathcal{C}_2 + \kappa_{23}\mathcal{C}_2}.$$

These equations imply that $\mathcal{C}_1 = \mathcal{C}_2 = 0$, which makes sense since the middle reaction allows the left two-cycle to drain completely into the right two-cycle. Thus any complex balanced solution of (5.21) only uses part of the network. ◁▷

To avoid this behavior, we make the following definition.

DEFINITION 5.14. A chemical reaction network (G, κ, Y) is *complex balanced* if it has a positive steady state solution x_* which is complex balanced, i.e., satisfies $A_\kappa\Phi(x_*) = 0$ with $x_* > 0$.

The key word in Definition 5.14 is *positive*, which implies that x_* and hence $\Phi(x_*) = (x_*^{y_1}, \ldots, x_*^{y_m})$ are both positive. Thus a complex balanced network has a complex balanced steady state that uses every complex of the network.

Strong Linkage Classes and Weak Reversibility. The network in (5.21) is connected as an ordinary graph but not as a directed graph. The two-cycles on the left and right are its *strong connected components* or *strong linkage classes*. The latter terminology is common in the CRNT literature. The strong linkage class on the left is *terminal* since nothing flows out of it. This explains the labels in (5.21).

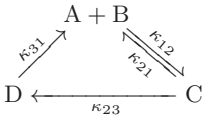
Given a chemical reaction network (G, κ, Y), it is straightforward to show that the following conditions are equivalent:

- Every connected component is strongly connected.
- Every complex lies in a terminal strong linkage class.
- For every reaction $C_i \longrightarrow C_j$, there is a sequence of reactions

$$C_j \longrightarrow C_k \longrightarrow \cdots \longrightarrow C_m \longrightarrow C_i.$$

DEFINITION 5.15. A chemical reaction network is *weakly reversible* if it satisfies any of the above conditions.

The network in Example 5.13 is not weakly reversible. In contrast, consider the network in Example 5.4:



The reaction $A + B \xrightarrow{\kappa_{12}} C$ is reversible via $C \xrightarrow{\kappa_{21}} A + B$, while $D \xrightarrow{\kappa_{31}} A + B$ is weakly reversible via $A + B \xrightarrow{\kappa_{12}} C \xrightarrow{\kappa_{23}} D$. The story for $C \xrightarrow{\kappa_{23}} D$ is similar, so the whole network is clearly weakly reversible.

Here is an important recent result about weakly reversible systems proved by Boros in 2019 [**37**]. When $x(0)$ is positive, it is customary to call $(x(0) + S) \cap \mathbb{R}^n_{>0}$ a *positive stoichiometric compatibility class.*

THEOREM 5.16. *If (G, κ, Y) is a weakly reversible chemical reaction network, then every positive stoichiometric compatibility class has a positive steady state.*

Dynamics of Complex Balanced Networks. The existence of a single positive complex balanced steady state has remarkable implications for the dynamics of the system. This is the Balanced Dynamics Theorem.

THEOREM 5.17 (Balanced Dynamics). *Let (G, κ, Y) be a complex balanced chemical reaction network. Then*:

(1) *The network is weakly reversible.*
(2) *Every positive stoichiometric compatibility class has a unique positive steady state.*
(3) *Every positive steady state is complex balanced and locally attracting within its stoichiometric compatibility class,*

PROOF. For (1), suppose that a positive steady state x_* satisfies $A_\kappa \Phi(x_*) = 0$. Since the reaction rates κ are positive, one can prove that $\ker(A_\kappa)$ is supported on the terminal strong linkage classes. (In [**148**, Lecture 4], Feinberg notes that this can be derived from the Perron-Frobenius Theorem. Self-contained proofs can be found in [**149**] and [**182**, Theorem 4.2].) The positivity of x_* implies that $\Phi(x_*) \in \ker(A_\kappa)$ is positive. It follows that every complex belongs to a terminal strong linkage class, proving weak reversibility.

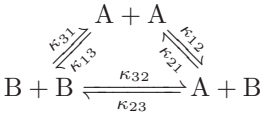
Proofs of (2) and (3) can be found in the 1972 paper of Horn and Jackson [**213**]. We recommend [**148**, Lectures 4 and 5] and [**182**, Theorem 6.4]. □

This powerful theorem tells us that the existence of a positive complex balanced steady state implies that all positive steady states are complex balanced. The name *Balanced Dynamics Theorem* is not standard and will change to *Toric Dynamics Theorem* in Section 5.2.

Theorem 5.17 also tells us that positive steady states are unique and locally attracting within their stoichiometric compatibility classes. Are they globally attracting? Jackson and Horn thought so in 1972, but later realized that their proof only implies that positive steady states are locally attracting. In 1974, Horn [**212**] formulated the *Global Attractor Conjecture*: in a complex balanced network, every positive steady state is globally attracting in its stoichiometric compatibility class. We will say more about this conjecture in Section 5.2.

One thing to keep in mind is that the existence of a positive complex balanced steady state may depend on the reaction rates. Here is an example.

EXAMPLE 5.18. Consider the network:



In Section 5.2, we will see that this network is complex balanced if and only if

$$(\kappa_{21}\kappa_{31}+\kappa_{23}\kappa_{31}+\kappa_{32}\kappa_{21})(\kappa_{13}\kappa_{23}+\kappa_{12}\kappa_{23}+\kappa_{21}\kappa_{13}) = (\kappa_{12}\kappa_{32}+\kappa_{13}\kappa_{32}+\kappa_{31}\kappa_{12})^2.$$

When this is satisfied, Theorem 5.17 describes the dynamics of the system. ◁▷

The Deficiency of a Network. The *deficiency* is a numerical invariant of a reaction network introduced by Feinberg in 1972 [**147**].

DEFINITION 5.19. The *deficiency* of (G, κ, Y) is

$$\delta = m - \ell - s,$$

where

- m = number of complexes = number of vertices of G.
- ℓ = number of connected components of the underlying ordinary graph.
- s = dimension of the stoichiometric subspace $S = \mathrm{Span}(y_j - y_i \mid i \to j \in E)$.

Feinberg proved $\delta \geq 0$ by showing that

$$\delta = m - \ell - s = \dim(\ker(Y) \cap \mathrm{im}(A_\kappa)). \tag{5.22}$$

In the weakly reversible case, we will give an algebro-geometric proof of $\delta \geq 0$ by interpreting δ as a codimension in Section 5.2.

The definition of δ involves only G and Y, yet by (5.22), it tells us something about A_κ. Here an easy example of how to exploit this.

LEMMA 5.20. *If (G, κ, Y) has deficiency zero, then for any positive κ, every positive steady state is complex balanced.*

PROOF. A positive steady state satisfies $Y A_\kappa \Phi(x_*) = 0$, so that

$$A_\kappa \Phi(x_*) \in \ker(Y) \cap \mathrm{im}(A_\kappa) = \{0\},$$

where the last equality follows from (5.22) and $\delta = 0$. Hence $A_\kappa \Phi(x_*) = 0$. □

The Deficiency Zero Theorem.. Here is the full version of the Deficiency Zero Theorem, proved by Horn in 1972 [**211**].

THEOREM 5.21 (Deficiency Zero). *A chemical reaction network is complex balanced for every set of positive reaction rates κ if and only if it has deficiency zero and is weakly reversible. Furthermore, in a weakly reversible network of deficiency zero, we have:*

(1) *For any positive rate constants κ and any positive stoichiometric compatibility class C, there is a unique positive steady state solution $x_* \in C$.*
(2) *The solution x_* is complex balanced and locally attracting within C.*

PROOF. If the network is weakly reversible and has deficiency zero, then:

- Weak reversibility and Theorem 5.16 imply that every positive stoichiometric compatibility class has a positive steady state solution.
- Deficiency zero and Lemma 5.20 imply that these positive steady states are complex balanced.

It follows that the network is complex balanced for any positive κ, and then parts (1) and (2) follow immediately from Theorem 5.17.

Going the other way, Theorem 5.17 tells us that complex balanced implies weakly reversible. Finally, we need to prove that $\delta = 0$ when the network complex balanced for all for positive κ. As we explain in Corollary 5.36, this follows from the Toric Deficiency Theorem (Theorem 5.35), to be discussed in Section 5.2. □

Other proofs of the Deficiency Zero Theorem can be found in [**148**] and [**182**]. Gatermann and Huber give an interesting proof in [**162**, Theorem 7.1].

Here are two simple examples.

EXAMPLE 5.22. Consider the following reaction network:

$$\begin{array}{ccc} A+C & \underset{\kappa_{21}}{\overset{\kappa_{12}}{\rightleftharpoons}} & D \\ & {\scriptstyle\kappa_{13}}\searrow \quad \nearrow{\scriptstyle\kappa_{32}} & \\ & B+E & \end{array} \qquad Y = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

This is weakly reversible. We compute the deficiency $\delta = m - \ell - s$:

- $m = 3$ (complexes) and $\ell = 1$ (connected components).
- $S = \mathrm{Span}(y_2 - y_1, y_1 - y_2, y_3 - y_1, y_2 - y_3) = \mathrm{Span}(y_3 - y_1, y_2 - y_3)$.

Thus $\delta = 3 - 1 - 2 = 0$, so Theorem 5.21 applies. ◁▷

EXAMPLE 5.23. In a similar way, one computes that $\delta = 1$ in Example 5.18, so the network cannot be complex balanced for all κ by Theorem 5.21. In Section 5.2 we will see that $\delta = 1$ implies that a single equation determines when κ is complex balanced, just as we claimed in Example 5.18. ◁▷

Final Comments. This section introduced chemical reaction networks and their associated systems of ODEs. We learned a little chemistry and saw a range of applications. There was also some nice mathematics, including a bit of algebraic geometry (e.g., positive steady states form a semialgebraic variety). In the next section, we will look more deeply into the combinatorics and algebraic geometry of chemical reaction networks, including the Matrix Tree Theorem and toric varieties.

5.2. Toric Dynamical Systems

In 2009, Craciun, Dickenstein, Shiu and Sturmfels [**113**] proposed the name *toric dynamical system* for differential equations of the form

$$\frac{dx}{dt} = Y A_\kappa \Phi(x)$$

that are complex balanced, i.e., there is a positive steady state x_* that satisfies the complex balancing equations $A_\kappa \Phi(x_*) = 0$. We use "toric dynamical system" for both the differential equations and the underlying reaction network (G, κ, Y).

In this setting, the Balanced Dynamics Theorem (Theorem 5.17) becomes the *Toric Dynamics Theorem*:

THEOREM 5.24 (Toric Dynamics). *Let (G, κ, Y) be a toric dynamical system. Then*:

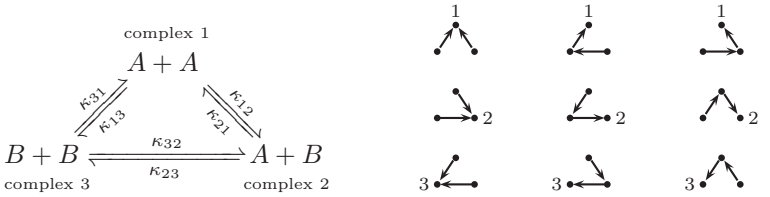
(1) *Every connected component of G is strongly connected (weak reversibility).*
(2) *Every positive stoichiometric compatibility class has a unique positive steady state.*
(3) *Every positive steady state is complex balanced and locally attracting within its stoichiometric compatibility class,*

The reason for changing to "toric dynamical system" and "Toric Dynamics Theorem" is the deep connection to toric varieties, which will emerge later in the section.

Example of a Toric Dynamical System. Before jumping into the general theory, let's study an example in detail.

EXAMPLE 5.25. Consider the following network:

(5.23)



For the moment, ignore the right-hand side of (5.23). For this network, we have

$$A_\kappa = \begin{bmatrix} -\kappa_{12} - \kappa_{13} & \kappa_{21} & \kappa_{31} \\ \kappa_{12} & -\kappa_{21} - \kappa_{23} & \kappa_{32} \\ \kappa_{13} & \kappa_{23} & -\kappa_{31} - \kappa_{32} \end{bmatrix},$$

and one computes that the kernel of A_κ is spanned by (K_1, K_2, K_3), where

$$\begin{aligned} K_1 &= \kappa_{21}\kappa_{31} + \kappa_{23}\kappa_{31} + \kappa_{32}\kappa_{21} \\ K_2 &= \kappa_{12}\kappa_{32} + \kappa_{13}\kappa_{32} + \kappa_{31}\kappa_{12} \\ K_3 &= \kappa_{13}\kappa_{23} + \kappa_{12}\kappa_{23} + \kappa_{21}\kappa_{13}. \end{aligned} \tag{5.24}$$

The K_i have a lovely combinatorial structure: each K_i is determined by the spanning trees with a unique sink at i, as shown on the right-hand side of (5.23). We will see below that this is a special case of the Matrix Tree Theorem.

As noted in Example 5.23, the deficiency is $\delta = m - \ell - s = 1$. Here are the details:

- $m = 3$ (complexes) and $\ell = 1$ (connected components).
- $Y = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix} \Longrightarrow S = \operatorname{Span}\big(y_i - y_j \mid i \neq j\big) = \operatorname{Span}\left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right) \Longrightarrow s = 1.$

Since this network is weakly reversible, the Deficiency Zero Theorem tells us that not all positive rate constants κ give a toric dynamical system. To find those that do, suppose that $x_* = (x_1, x_2)$ is positive and satisfies $A_\kappa \Phi(x_*) = 0$. Using the

above description of $\ker(A_\kappa)$, we obtain

$$\Phi(x_*) \in \ker(A_\kappa) \Longrightarrow \begin{bmatrix} x_1^2 \\ x_1x_2 \\ x_2^2 \end{bmatrix} = \lambda \begin{bmatrix} K_1 \\ K_2 \\ K_3 \end{bmatrix} \Longrightarrow K_1K_3 = K_2^2 \tag{5.25}$$

since $\lambda \neq 0$ by the positivity of x_*. Later in the section we will use the Toric Moduli Theorem (Theorem 5.34) to prove the converse, so that we have a toric dynamical system if and only if $K_1K_3 = K_2^2$, as claimed in Example 5.18.

To complete the example, let's look more deeply into the algebra of what we just did. From (5.25), we get the system of equations

$$x_1^2 = \lambda K_1,\ x_1x_2 = \lambda K_2,\ x_2^2 = \lambda K_3$$

in variables $x_1, x_2, K_1, K_2, K_3, \lambda$, Eliminating λ, we get the ideal

$$\langle K_2x_1^2 - K_1x_1x_2, K_3x_1x_2 - K_2x_2^2, K_3x_1^2 - K_1x_2^2, K_3x_1^3 - K_2x_1^2x_2, x_1^3(K_1K_3 - K_2^2)\rangle.$$

Note that the first three generators are $K_ix^{y_j} - K_jx^{y_i}$ for $i > j$, which suggests that something systematic is going on here.

Since we want $x_* = (x_1, x_2)$ to be positive, they should both be nonzero. Hence it makes sense to saturate with respect to x_1x_2. Doing so gives the ideal

$$T_G = \langle K_2x_1 - K_1x_2, K_3x_1 - K_2x_2, K_1K_3 - K_2^2\rangle.$$

We will see below that T_G is a toric ideal. Then the equation $K_1K_3 = K_2^2$ we first saw in Example 5.18 arises naturally from the elimination ideal

$$M_G = T_G \cap \mathbb{Q}[K_1, K_2, K_3] = \langle K_1K_3 - K_2^2\rangle.$$

This is another toric ideal. ◁▷

Changing Variables. In terms of κ, the equation $K_1K_3 = K_2^2$ becomes

$$(\kappa_{21}\kappa_{31}+\kappa_{23}\kappa_{31}+\kappa_{32}\kappa_{21})(\kappa_{13}\kappa_{23}+\kappa_{12}\kappa_{23}+\kappa_{21}\kappa_{13}) = (\kappa_{12}\kappa_{32}+\kappa_{13}\kappa_{32}+\kappa_{31}\kappa_{12})^2,$$

which defines a complicated variety. The toric nature of this example emerges only after we make the change of variables given by (5.24).

The phenomenon of "toric emergence" happens in other contexts and in particular occurs in an example from algebraic statistics that appeared in Chapter 2.

Example 5.26. In Example 2.17, we considered a phylogenetic model from [**237**] and in (2.26) observed that the probabilities p_{ijk}, $i, j, k \in \{0, 1\}$, satisfy the obvious linear equation $p_{000}+\cdots+p_{111} = 1$ together with three quadratic equations and many inequalities. Let's focus on the quadratic polynomials, one of which is

$$p_{001}p_{010}+p_{001}p_{100}-p_{000}p_{011}-p_{000}p_{101}+p_{100}p_{111}-p_{101}p_{110}+p_{010}p_{111}-p_{011}p_{110}.$$

The other two are similar. These polynomials are not overly complicated but, as explained by Sturmfels and Sullivant in [**337**, Example 3], they become even nicer via the change of coordinates

$$p_{ijk} = \sum_{r,s,t=0}^{1} (-1)^{ir+js+kt} q_{rst}.$$

This can be regarded as a finite Fourier transform. In the new coordinates, the three quadratic polynomials transform to

$$q_{001}q_{110} - q_{000}q_{111},\ q_{010}q_{101} - q_{000}q_{111},\ q_{100}q_{011} - q_{000}q_{111}.$$

These are not only simpler but also define a toric variety! ◁▷

In the situation of Example 5.25, changing from $\kappa_{12}, \kappa_{21}, \kappa_{34}, \kappa_{43}, \kappa_{45}, \kappa_{54}$ to K_1, K_2, K_3 gives the toric equation $K_1 K_3 = K_2^2$. To do this change of variables in general, we need the Matrix Tree Theorem.

The Matrix Tree Theorem. Toric dynamical systems are weakly reversible by Theorem 5.24. For this reason, in the rest of the section we will only consider networks (G, κ, Y) where G is weakly reversible. Hence every connected component of G is strongly connected. Let $G_1, \dots, G_\ell$ be the connected components of G. We also assume that the the vertices of G are the set $[m] = \{1, \dots, m\}$.

In this situation, for $i \in [m]$, define

$$K_i = \sum_{T \downarrow i} \kappa^T \tag{5.26}$$

where:

- The sum is over all maximal subtrees T of G with a unique sink at i. We denote this by writing $T \downarrow i$.
- κ^T is the product of rate constants corresponding to edges of T.

When the vertex i lies in the connected component G_ν, a maximal subtree of G containing i is simply a spanning tree of G_ν. So the above sum is over all spanning trees of G_ν that have a unique sink at i. We have already seen an example of K_i in (5.23) and (5.24).

Recall that $A_\kappa = -L_\kappa$, where L_κ is the graph Laplacian defined in (5.12). Using the K_i, we can describe the kernel of A_κ as follows.

THEOREM 5.27 (Matrix Tree). *If G is a weakly reversible directed graph with connected components $G_1, \dots, G_\ell$ and positive edge weights $\kappa = (\kappa_{i \to j})$, then:*

(1) $\dim(\ker(A_\kappa)) = \ell$.
(2) *A basis of* $\ker(A_\kappa) \subseteq \mathbb{R}^m$ *is given by*

$$v_\nu = \sum_{i \text{ vertex of } G_\nu} K_i e_i, \quad \nu = 1, \dots, \ell,$$

where K_i is defined in (5.26) *and $\{e_1, \dots, e_m\}$ is the standard basis of $\mathbb{R}^m$.*
(3) *If G is connected and B_κ is the $(m-1) \times m$ matrix obtained by removing the last row of A_κ, then $K_1, \dots, K_m$ are the signed maximal minors of B_κ.*

PROOF. For (1) and (2), observe that after reordering the vertices, $A_\kappa = A_{G,\kappa}$ becomes a block matrix whose diagonal blocks are A_{G_ν, κ_ν}, $\nu = 1, \dots, \ell$, with 0's elsewhere. So we may assume that G is connected and thus strongly connected. In this situation, the trees T in the definition of K_i are spanning trees of G with a unique sink at i, and (1) and (2) reduce to the assertion that $(K_1, \dots, K_m)$ spans $\ker(A_\kappa)$. This follows from Lemma 1 and Theorem 1 in [**183**, 9.13.4], which includes a nice discussion of the Matrix Tree Theorem and references to the proof. For another proof, see [**162**, Theorem 3.2].

For (3), let B_κ^i be obtained from B_κ by removing column i. We need to prove that $K_i = (-1)^{i-1} \det(B_\kappa^i)$. Theorem 3 of [**123**] implies that $K_i = \det(L_\kappa^{i,i})$, where the superscript "i, i" denotes the matrix obtained by deleting row i and column i of L_κ. Since the column sums of L_κ are 0, it follows that

$$\begin{aligned} K_i &= \det(L_\kappa^{i,i}) = (-1)^{i-m} \det(L_\kappa^{i,m}) \\ &= (-1)^{i-m}(-1)^{m-1} \det(A_\kappa^{i,m}) = (-1)^{i-1} \det(B_\kappa^i), \end{aligned}$$

since $L_\kappa = -A_\kappa$ and B_κ is obtained from A_κ by removing row m (the last row). □

EXAMPLE 5.28. Removing the last row of the matrix A_κ in Example 5.25 gives

$$B_\kappa = \begin{bmatrix} -\kappa_{12}-\kappa_{13} & \kappa_{21} & \kappa_{31} \\ \kappa_{12} & -\kappa_{21}-\kappa_{23} & \kappa_{32} \end{bmatrix},$$

and one computes without difficulty that the signed maximal minors are precisely K_1, K_2, K_3 from (5.24). ◁▷

Ideals, Varieties and Chemistry. This is the title of Section 2 of [**113**], which introduces the algebraic language used to describe toric dynamical systems. Given (G, κ, Y) with $G = ([m], E)$ weakly reversible, we have as usual the ODE system

$$\frac{d}{dt}x = Y A_\kappa \Phi(x)$$

for concentrations $x = (x_1, \ldots, x_n)^t$. We regard the rate constants κ_{ij} for $i \to j \in E$ as variables, giving the polynomial ring $\mathbb{Q}[x, \kappa] = \mathbb{Q}[x_1, \ldots, x_n, \kappa_{ij} \mid i \to j \in E]$.

For a toric dynamical system, we seek positive solutions of $A_\kappa(\Phi(x)) = 0$. Algebraically, the closest we can come to "$x_i > 0$" is "$x_i \neq 0$". As in Example 5.25, this means saturating with respect $x_1 \cdots x_n$. Hence we get the ideal

$$C_G = \langle A_\kappa \Phi(x) \rangle : (x_1 \cdots x_n)^\infty \subseteq \mathbb{Q}[x, \kappa].$$

If $\mathbf{V}_{>0}(\cdot)$ denotes solutions whose coordinates lie in $\mathbb{R}_{>0}$, then the semialgebraic variety

$$\mathbf{V}_{>0}(C_G) \subseteq \mathbb{R}^n_{>0} \times \mathbb{R}^E_{>0}$$

has the following lovely interpretation: if $(x_*, \kappa_*) \in \mathbb{R}^n_{>0} \times \mathbb{R}^E_{>0}$, then

$$(x_*, \kappa_*) \in \mathbf{V}_{>0}(C_G) \iff \begin{array}{l} (G, \kappa_*, Y) \text{ is a toric dynamical system and} \\ x_* \text{ is a positive complex balanced steady state.} \end{array} \tag{5.27}$$

In other words, $(x_*, \kappa_*) \in \mathbf{V}_{>0}(C_G)$ means that x_* is a "witness" or "certificate" proving that (G, κ_*, Y) is a toric dynamical system.

Following [**113**], we next describe some interesting elements of C_G. Suppose that $(x_*, \kappa_*) \in \mathbf{V}_{>0}(C_G)$. Then let $K_* = (K_{*1}, \ldots, K_{*m}) \in \mathbb{R}^m_{>0}$ be obtained by evaluating the K_i at κ_*. Since $A_{\kappa_*}(\Phi(x_*)) = 0$, we have

$$\Phi(x_*) = \sum_{\nu=1}^{\ell} \lambda_\nu \sum_{k \text{ vertex of } G_\nu} K_{*k} e_k$$

by the Matrix Tree Theorem. In particular, if i, j are vertices of the same connected component G_ν of G, then

$$\left.\begin{aligned} x_*^{y_i} &= \lambda_\nu K_{*i} \\ x_*^{y_j} &= \lambda_\nu K_{*j} \end{aligned}\right\} \Longrightarrow K_{*i} x_*^{y_j} - K_{*j} x_*^{y_i} = 0.$$

Thus the polynomial $K_i x^{y_j} - K_j x^{y_i}$ vanishes on $\mathbf{V}_{>0}(C_G)$. The intrinsic reasoning for this vanishing is given by the following lemma.

LEMMA 5.29. *$K_i x^{y_j} - K_j x^{y_i} \in C_G$ when i, j are vertices in the same connected component of G.*

PROOF. Following the proof of Theorem 5.27, we can assume G is connected. If we set $A_k(\Phi(x)) = (u_1, \ldots, u_m)^t$, then $u_1, \ldots, u_m \in C_G$.

For simplicity, we will prove $K_1 x^{y_2} - K_2 x^{y_1} \in C_G$. As in Theorem 5.27, we remove the last row of A_κ to get an $(m-1) \times m$ matrix B_κ. Thus $B_\kappa(\Phi(x)) =$

$(u_1, \ldots, u_{m-1})^t$. Let $\mathbf{w}_1, \ldots, \mathbf{w}_m$ be columns of B_κ, i.e., $B_\kappa = [\mathbf{w}_1 \cdots \mathbf{w}_m]$. Given any row vector $D = (D_1, \ldots, D_{m-1})$, one computes that

$$\begin{aligned} D_1u_1 + \cdots + D_{m-1}u_{m-1} &= D(B_\kappa\Phi(x)) = (DB_\kappa)\Phi(x) \\ &= (D \cdot \mathbf{w}_1)x^{y_1} + \cdots + (D \cdot \mathbf{w}_m)x^{y_m}. \end{aligned}$$

Now let $D_1, \ldots, D_{m-1}$ be the signed maximal minors of the $(m-2) \times (m-1)$ matrix $[\mathbf{w}_3 \cdots \mathbf{w}_m]$ obtained by deleting first two columns of B_κ. Expanding by minors along the first column shows that

$$\det[\mathbf{w}\ \mathbf{w}_3 \cdots \mathbf{w}_m] = D \cdot \mathbf{w}$$

for any column vector $\mathbf{w}$. Then the above formula simplifies to

$$\begin{aligned} D_1u_1 + \cdots + D_{m-1}u_{m-1} &= \det[\mathbf{w}_1\ \mathbf{w}_3 \cdots \mathbf{w}_m]x^{y_1} + \det[\mathbf{w}_2\ \mathbf{w}_3 \cdots \mathbf{w}_m]x^{y_2} \\ &= \det(B_\kappa^2)x^{y_1} + \det(B_\kappa^1)x^{y_2} = -K_2x^{y_1} + K_1x^{y_2}, \end{aligned}$$

where the last line uses $K_i = (-1)^{i-1}\det(B_\kappa^i)$ from the proof of Theorem 5.27. □

A Change of Variables and a New Ideal. Once we regard the rate constants κ as variables, the K_i give the subring

$$\mathbb{Q}[K] = \mathbb{Q}[K_1, \ldots, K_m] \subseteq \mathbb{Q}[\kappa].$$

By Lemma 5 of [**113**], the K_i are algebraically independent over $\mathbb{Q}$. Hence we can regard $\mathbb{Q}[K]$ as a polynomial ring. This leads to the ideal

$$T_G = \langle K_i x^{y_j} - K_j x^{y_i} \mid i, j \text{ in same connected component of } G \rangle : (x_1 \cdots x_n)^\infty$$

in the polynomial ring $\mathbb{Q}[x, K]$. This is a toric ideal (prime and generated by binomials), as proved in [**113**, Proposition 6]. A proof using multi-Rees algebras can be found in [**107**, Theorem 3.3].

The following basic result explains the relevance of the toric ideal T_G.

PROPOSITION 5.30. *Given $(x_*, \kappa_*) \in \mathbb{R}^n_{>0} \times \mathbb{R}^E_{>0}$, the following are equivalent*:

(1) *(G, κ_*, Y) is a toric dynamical system and x_* is a positive complex balanced steady state.*
(2) $(x_*, \kappa_*) \in \mathbf{V}_{>0}(C_G)$.
(3) $(x_*, K_*) \in \mathbf{V}_{>0}(T_G)$.

PROOF. The equivalance (1) ⇔ (2) is (5.27), and (2) ⇒ (3) follows from the discussion leading up to Lemma 5.29. For (3) ⇒ (1), note that K_{*i} and $x_*^{y_i}$ are positive for all i. By the definition of T_G, the ratio $x_*^{y_i}/K_i$ depends only on the connected component of G containing i. If we set

$$\lambda_\nu = \frac{x_*^{y_i}}{K_{*i}},\ i \text{ any vertex of } G_\nu,$$

then

$$\Phi(x_*) = \sum_{\nu=1}^{\ell} \sum_{k \text{ vertex of } G_\nu} x_*^{y_k} e_k = \sum_{\nu=1}^{\ell} \lambda_\nu \sum_{k \text{ vertex of } G_\nu} K_{*k} e_k,$$

which is in $\ker(A_{\kappa_*})$ by the Matrix Tree Theorem. Thus x_* is complex balanced for the rate constants κ_*, proving that (G, κ_*, Y) is a toric dynamical system. □

Using the change of coordinates from κ to K, we now have toric ideal that tells us exactly when we have a toric dynamical system with a witness.

EXAMPLE 5.31. In the situation of Example 5.25, the ideal $T_G \subseteq \mathbb{Q}[x, K] = \mathbb{Q}[x_1, x_2, K_1, K_2, K_3]$ is

$$T_G = \langle K_2x_1^2 - K_1x_1x_2, K_3x_1^2 - K_1x_2^2, K_3x_1x_2 - K_2x_2^2\rangle : (x_1x_2x_3)^\infty$$

Using the method explained in Theorem 14 of [**104**, Chapter 4, §4], we obtain

$$T_G = \langle K_2x_1 - K_1x_2, K_3x_1 - K_2x_2, K_1K_3 - K_2^2\rangle,$$

exactly as claimed in Example 5.25. ◁▷

Moduli of Toric Dynamical Systems. The next step is to get rid of the witness x_*. In other words, what conditions must κ_* satisfy in order for x_* to exist?

EXAMPLE 5.32. Let's see what happens in Example 5.31. Suppose κ_* satisfies $K_{*1}K_{*3} = K_{*2}^2$. We know from Example 5.25 that the stoichiometric subspace S is defined by $x_1 + x_2 = 0$. Given $c > 0$, the point

$$x_* = c\Big(\frac{K_{*1}}{K_{*1} + K_{*2}}, \frac{K_{*2}}{K_{*1} + K_{*2}}\Big) \in \mathbb{R}^2_{>0}$$

lies in the stoichiometric compatibility class $x_1+x_2 = c$. More importantly, (x_*, K_*) clearly satisfies $K_{*2}x_{*1} - K_{*1}x_{*2} = 0$, and then $K_{*3}x_{*1} - K_{*2}x_{*2} = 0$ follows from $K_{*1}K_{*3} = K_{*2}^2$. Thus $(x_*, K_*) \in \mathbf{V}_{>0}(T_G)$, so (G, κ_*, Y) is complex balanced. This proves the claim about complex balancing made in Example 5.18. ◁▷

This example is nice, but we want theorems, not computations. The idea is to use elimination theory. Given $T_G \subseteq \mathbb{Q}[x, K]$, we define the *toric moduli ideal*

$$M_G = T_G \cap \mathbb{Q}[K] \subseteq \mathbb{Q}[K].$$

Note that M_G is an elimination ideal.

LEMMA 5.33. *M_G is a toric ideal.*

PROOF. We noted above that T_G is a toric ideal. Thus M_G is prime since T_G is prime. Applying the Buchberger algorithm to the binomial generators of T_G shows that T_G has a binomial lex Gröbner basis. Hence M_G has binomial generators by the Elimination Theorem. □

The key property of M_G is that it defines the moduli space of toric dynamical systems for G and Y.

THEOREM 5.34 (Toric Moduli). *Assume G is weakly reversible. For $\kappa_* \in \mathbb{R}^E_{>0}$, (G, κ_*, Y) is a toric dynamical system if and only if $K_* \in \mathbf{V}_{>0}(M_G)$.*

PROOF. This is Theorem 7 from [**113**]. The intersection $M_G = T_G \cap \mathbb{Q}[K]$ gives a commutative diagram of inclusions and projections:

$$\begin{array}{ccc} \mathbf{V}_{>0}(T_G) & \hookrightarrow & \mathbb{R}^n_{>0} \times \mathbb{R}^m_{>0} \\ \downarrow \pi & & \downarrow \\ \mathbf{V}_{>0}(M_G) & \hookrightarrow & \mathbb{R}^m_{>0}, \end{array}$$

with an analogous diagram over $\mathbb{C}$.

If (G, κ_*, Y) is a toric dynamical system, then by Lemma 5.30, there is x_* such that $(x_*, K_*) \in \mathbf{V}_{>0}(T_G)$. Thus $K_* = \pi(x_*, K_*) \in \mathbf{V}_{>0}(M_G)$. For the other direction, assume κ_* gives $K_* \in \mathbf{V}_{>0}(M_G)$. We need to prove the existence of a witness x_* with $(x_*, K_*) \in \mathbf{V}_{>0}(T_G)$. If we work over the complex numbers, the

Closure Theorem (Theorem 3 of [**104**, Chapter 3, §2]) implies that π has Zariski dense image, i.e.,

$$\overline{\pi(\mathbf{V}_{\mathbb{C}}(T_G))} = \mathbf{V}_{\mathbb{C}}(M_G). \tag{5.28}$$

Since T_G and $M_G = T_G \cap \mathbb{Q}[K]$ are toric ideals, the projection map

$$\pi : \mathbf{V}_{\mathbb{C}}(T_G) \to \mathbf{V}_{\mathbb{C}}(M_G)$$

is a toric morphism, which induces a group homomorphism of the associated complex tori. The image of this map is Zariski dense by (5.28) and is a subtorus by the theory of algebraic groups. Thus the map on tori is surjective.

Since $K_* \in \mathbf{V}_{>0}(M_G)$ lies in the torus of $\mathbf{V}_{\mathbb{C}}(M_G)$, it follows K_* is the image of a point in the torus of $\mathbf{V}_{\mathbb{C}}(T_G)$, which we write as $(z_1, \ldots, z_n, K_*) \in \mathbf{V}_{\mathbb{C}}(T_G)$ with $z_1, \ldots, z_n \in \mathbb{C}^*$. We have a complex witness! Since K_* is positive and T_G is a binomial ideal, it follows that $x_* = (|z_1|, \ldots, |z_n|)$ satisfies $(x_*, K_*) \in \mathbf{V}_{>0}(T_G)$. Thus x_* is the desired witness. □

The proof of Theorem 5.34 is a lovely combination of classical elimination theory (which works over $\mathbb{C}$) and toric geometry (which enables us to translate from $\mathbb{C}$ to $\mathbb{R}_{>0}$). The statement of the theorem reminds us that we always assume that G is weakly reversible, since otherwise (G, κ_*, Y) is never complex balanced.

Our final result from [**113**] concerns the deficiency.

THEOREM 5.35 (Toric Deficiency). *If G is weakly reversible, then the deficiency δ is the codimension of $\mathbf{V}_{\mathbb{C}}(M_G) \subseteq \mathbb{C}^m$.*

We refer the reader to [**113**, Theorem 9] for the proof. Here is a nice corollary.

COROLLARY 5.36. *If (G, κ, Y) is complex balanced for all positive κ, then it has deficiency zero.*

PROOF. The polynomials $K_1, \ldots, K_m \in \mathbb{Q}[\kappa]$ define a map $\mathcal{K} : \mathbb{C}^E \to \mathbb{C}^m$. Our hypothesis and Theorem 5.34 imply $\mathcal{K}(\mathbb{R}^E_{>0}) \subseteq \mathbf{V}_{>0}(M_G)$, so that we have inclusions

$$\mathcal{K}(\mathbb{R}^E_{>0}) \subseteq \mathbf{V}_{>0}(M_G) \subseteq \mathbf{V}_{\mathbb{C}}(M_G) \subseteq \mathbb{C}^m.$$

Since $\mathbb{R}^E_{>0}$ is Zariski dense in $\mathbb{C}^E$, we obtain

$$\mathcal{K}(\mathbb{C}^E) = \mathcal{K}\big(\overline{\mathbb{R}^E_{>0}}\big) \subseteq \overline{\mathcal{K}(\mathbb{R}^E_{>0})} \subseteq \mathbf{V}_{\mathbb{C}}(M_G) \subseteq \mathbb{C}^m.$$

But $\mathcal{K}$ is dominating since $K_1, \ldots, K_m$ are algebraically independent over $\mathbb{Q}$. This forces $\mathbf{V}_{\mathbb{C}}(M_G) = \mathbb{C}^m$, hence $\delta = 0$. □

This corollary can also be proved using the moment map and the Cayley polytope of G. See the discussion following Theorem 9 in [**113**].

The Work of Karin Gatermann. Gatermann wrote four influential papers that introduced toric methods into the study of chemical reaction networks:

[**161**] *Counting stable solutions of sparse polynomial systems in chemistry*, 2001.

[**162**] *A family of sparse polynomial systems arising in chemical reaction systems* (with B. Huber), 2002.

[**163**] *Bernstein's second theorem and Viro's method for sparse polynomial systems in chemistry* (with M. Wolfrum), 2005.

[**164**] *Toric ideals and graph theory to analyze Hopf bifurcations in mass action systems* (with M. Eiswirth and A. Sensse), 2005.

These papers inspired the authors of [**113**] to carry forward her ideas after her early death. In a MathSciNet review of her last paper [**164**], Maurice Rojas wrote:

> Karin Gatermann was a pioneer in computational algebra and extending its applications to differential equations, dynamical systems, and chemistry. Tragically, she succumbed to cancer (at a young age) on New Year's Day 2005. She is sorely missed by her colleagues. [**295**]

The introduction to Gatermann's 2002 paper with Huber provides good sense of what she was hoping to accomplish:

> The study of positive solutions by chemists and applied mathematicians is completely independent of the literature on sparse polynomial systems in algebraic geometry. ... The aim of this paper is to build a bridge between the two areas, the applied mathematics literature and the algebraic literature on sparse polynomial systems. [**162**]

As an example of the success of this bridge, we note that the proof of Theorem 5.35 in [**113**] is a variation of the Cayley trick argument that Gatermann and Wolfrum use in [**163**] in their study of complex balanced systems.

To better understand of the depth of Gatermann's contribution, we will look at an example proposed by Edelstein in 1970 [**137**]. We will explore this example from three perspectives:

1979: Classical tools of CRNT, based on Feinberg [**147**].
1989: Symbolic computation, based on Melenk, Möller and Neun [**269**].
2005: Toric methods, based on Gatermann and Wolfrum [**163**].

Edelstein 1970 and Feinberg 1979. In 1970, Edelstein [**137**] proposed a biochemical model

$$\mathrm{A} + \mathrm{L} \rightleftharpoons 2\,\mathrm{A} \qquad \mathrm{A} + \mathrm{B} \rightleftharpoons \mathrm{C} \rightleftharpoons \mathrm{B} + \mathrm{M} \tag{5.29}$$

where L and M are kept at fixed concentrations. Edelstein notes that this system can have multiple steady states and is analytically simple. He also comments that it has "not been biochemically demonstrated."

In his 1979 lectures, Feinberg [**148**, Lecture 2] notes that it is easier to work with

$$\mathrm{A} \underset{\kappa_{21}}{\overset{\kappa_{12}}{\rightleftharpoons}} 2\,\mathrm{A} \qquad \mathrm{A} + \mathrm{B} \underset{\kappa_{43}}{\overset{\kappa_{34}}{\rightleftharpoons}} \mathrm{C} \underset{\kappa_{54}}{\overset{\kappa_{45}}{\rightleftharpoons}} \mathrm{B} \tag{5.30}$$

where the concentrations of L and M have been absorbed into κ_{12} and κ_{54}. Thus (5.30) is a system which models the behavior of (5.29). In practice (including this book), (5.30) is often attributed to Edelstein, though in reality it is Feinberg's modification of Edelstein's original network (5.29).

The network (5.30) is an example of what Feinberg calls a "funny" reaction network. At the beginning of his discussion of these networks, he says

> ... we shall admit for consideration networks containing peculiar reactions like $\mathrm{A} \longrightarrow 2\,\mathrm{A}$ or $0 \longrightarrow \mathrm{A}$ (zero reacts to A) which, at first glance, appear to be incompatible with conservation of matter. [**148**, 2-20]

Such reactions enable mass-action networks to model stirred tanks (5.20), which are "open" systems because reactants are added and products are removed.

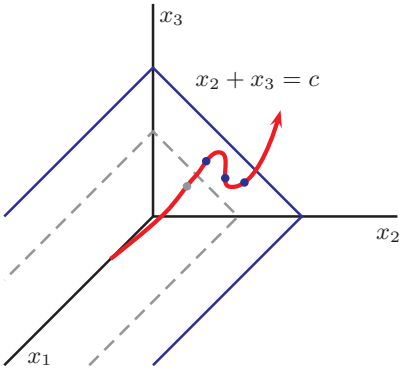
One reason Edelstein and Feinberg like this example is that it has different dynamical behavior for different values of the reaction rates κ. We will see below that when $\kappa_{21}\kappa_{43}\kappa_{54} = \kappa_{12}\kappa_{34}\kappa_{45}$, (5.30) is a toric dynamical system with one

steady state in each positive stoichiometric compatibility class. But for the reaction rates

$$\text{A} \underset{1}{\overset{8.5}{\rightleftharpoons}} 2\,\text{A} \qquad \text{A}+\text{B} \underset{1}{\overset{1}{\rightleftharpoons}} \text{C} \underset{0.2}{\overset{1}{\rightleftharpoons}} \text{B} \tag{5.31}$$

from [**148**, 3.1], the behavior is more varied. The number of steady states in a given stoichiometric compatibility class can vary from 1 to 3, as shown here:

(5.32)



We will say more about this picture below.

CRNT Analysis of Edelstein's Example. For (5.30), we have

$$Y = \begin{bmatrix} 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \qquad S = \text{Span}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \right), \tag{5.33}$$

so for $c > 0$, the stoichiometric compatibility class is the intersection of $x_2 + x_3 = c$ with the positive orthant, as shown in (5.32). Note that this is an example of an unbounded stoichiometric compatibility class. We also have the ODEs

$$\begin{aligned} \frac{dx_1}{dt} &= \kappa_{12}x_1 - \kappa_{21}x_1^2 - \kappa_{34}x_1x_2 + \kappa_{43}x_3 \\ \frac{dx_2}{dt} &= -\kappa_{34}x_1x_2 + \kappa_{43}x_3 + \kappa_{45}x_3 - \kappa_{54}x_2 \\ \frac{dx_3}{dt} &= \kappa_{34}x_1x_2 - \kappa_{43}x_3 - \kappa_{45}x_3 + \kappa_{54}x_2. \end{aligned} \tag{5.34}$$

Note that (5.30) is weakly reversible. Our methods make it easy to determine when we get a toric dynamical system. By looking the maximal trees in (5.30) with a unique sink at complex 1, 2, 3, 4 or 5, one gets

$$K_1 = \kappa_{21},\ K_2 = \kappa_{12},\ K_3 = \kappa_{43}\kappa_{54},\ K_4 = \kappa_{34}\kappa_{54},\ K_5 = \kappa_{34}\kappa_{45}.$$

Now we compute some ideals:

$$\begin{aligned} T_G &= \langle K_1x_1^2 - K_2x_1, K_3x_3 - K_4x_1x_2, K_3x_2 - K_5x_1x_2, K_4x_2 - K_5x_3 \rangle : (x_1x_2x_3)^\infty \\ &= \langle K_1K_3 - K_2K_5, K_4x_2 - K_5x_3, K_5x_1 - K_3, K_1x_1 - K_2 \rangle \\ M_G &= T_G \cap \mathbb{Q}[K_1, K_2, K_3, K_4, K_5] = \langle K_1K_3 - K_2K_5 \rangle \end{aligned}$$

By Theorem 5.34, it follows that (5.34) is a toric dynamical system if and only if $K_1K_3 = K_2K_5$, or equivalently, $\kappa_{21}\kappa_{43}\kappa_{54} = \kappa_{12}\kappa_{34}\kappa_{45}$.

Note also that the reaction rates that give a toric dynamical system are defined by a single equation and hence have codimension one. Since the deficiency of (5.30) is $\delta = m - \ell - s = 5 - 2 - 2 = 1$, this is consistent with Theorem 5.35.

Melenk, Möller and Neun 1989. In their paper *Symbolic solution of large stationary chemical kinetics problems* [**269**], Melenk, Möller and Neun apply Gröbner basis methods to chemical reaction networks. They study Edelstein's example (5.30), though an unfortunate typographical error mars their analysis (the formula for p_3 in [**269**, 5.2] should have $c_1^2\beta$, not $c_1\beta$).

We instead do our own Gröbner basis analysis of (5.30). In the stoichiometric compatibility class $x_2 + x_3 = c$, (5.34) tells us that the steady states are defined by

$$\begin{aligned} \kappa_{12}x_1 + \kappa_{43}x_3 - \kappa_{21}x_1^2 - \kappa_{34}x_1x_2 &= 0 \\ \kappa_{43}x_3 + \kappa_{45}x_3 - \kappa_{34}x_1x_2 - \kappa_{54}x_2 &= 0 \\ x_2 + x_3 &= c. \end{aligned}$$

(We ignore the third line of (5.34) since it is the negative of the second.). We want κ and c to be arbitrary, which means computing a lex Gröbner basis for $x_3 > x_2 > x_1$ over $\mathbb{Q}(\kappa, c)$. This gives the polynomials

$$\begin{aligned} g_1 &= (\kappa_{45} + \kappa_{54})x_2 - \kappa_{21}x_1^2 + \kappa_{12}x_1 - c\kappa_{45} \\ g_2 &= (\kappa_{45} + \kappa_{54})x_3 + \kappa_{21}x_1^2 - \kappa_{12}x_1 - c\kappa_{54} \\ g_3 &= k_{21}k_{34}x_1^3 + (-k_{12}k_{34} + k_{21}k_{43} + k_{21}k_{45} + k_{21}k_{54})x_1^2 \\ &\quad + (-k_{12}k_{43} - k_{12}k_{45} + ck_{34}k_{45} - k_{12}k_{54})x_1 - ck_{43}k_{54}. \end{aligned}$$

Since g_3 is cubic in x_1, there are 1, 2 or 3 values for x_1. The polynomials g_1 and g_2 enable us to solve for x_2 and x_3 in terms of x_1. It follows that there are potentially 1, 2 or 3 steady states in each stoichiometric compatibility class.

The discriminant D of g_3 is a (large) cubic polynomial in c. There are three real roots when $D > 0$ and one real root when $D < 0$. The coefficient of c^3 in D is $-4\kappa_{21}\kappa_{34}^4\kappa_{45}^3 < 0$, hence

$$c \gg 0 \implies D < 0 \implies g_3 \text{ has only one real root.} \tag{5.35}$$

From x_1, we get x_2 and x_3 using g_1 and g_2. It follows that the plane $x_2 + x_3 = c$ contains exactly one real steady state when $c \gg 0$. But it is positive? This is where Theorem 5.16 comes into play, for it guarantees the existence of at least one positive steady state satisfying $x_2 + x_3 = c$ when $c > 0$. Hence this positive steady state is unique when $c \gg 0$, exactly as predicted in the picture (5.32).

For the specific reaction rates in (5.31), the Gröbner basis becomes

$$\begin{aligned} g_1 &= 1.2x_2 - x_1^2 + 8.5x_1 - c \\ g_2 &= 1.2x_3 + x_1^2 - 8.5x_1 - .2c \\ g_3 &= x_1^3 - 6.3x_1^2 + (c - 18.7)x_1 - .2c. \end{aligned}$$

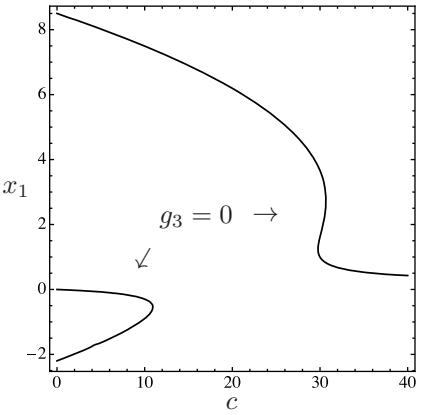
Thinking of g_3 as a polynomial in x_1 and c, we get the picture (5.36) on the next page. The discriminant D of g_3 has roots $c = 10.951, 29.777, 30.695$ in this case,

which when combined with the picture implies

$$\begin{aligned} c < 10.951 &\implies g_3 \text{ has three real roots, two of which are negative} \\ 10.951 < c < 29.777 &\implies g_3 \text{ has one real root, which is positive} \\ 29.777 < c < 30.695 &\implies g_3 \text{ has three real roots, all positive} \\ 30.695 < c &\implies g_3 \text{ has one real root, which is positive.} \end{aligned}$$

You can see this clearly in the picture (5.36).

(5.36)



A cylindrical algebraic decomposition of the system

$$g_1 = g_2 = g_3 = 0,\ c > 0,\ x_1 > 0,\ x_2 > 0,\ x_3 > 0$$

is easily done in MATHEMATICA using the `CylindricalDecomposition` command and reveals that when x_1 and c are positive, so are x_2 and x_3. Hence we recover the full picture shown in (5.32), and we now know that multistationarity occurs in a narrow range of stoichiometric compatibility classes around $c = 30$.

Gatermann and Wolfrum 2005. Before we can discuss what Gatermann and Wolfrum say about Edelstein's example in [**163**], we need to explain the framework of their paper. One difference is that unlike [**113**] and [**163**], we find it convenient to regard $\kappa = (\kappa_{ij})$ as an element of $\mathbb{R}^r_{>0}$ rather than a vector of variables.

Let's begin with the structure of the matrix A_κ. Suppose the directed graph G has m vertices (the complexes) and r edges (the reactions). If we enumerate the edges, then G has an $m \times r$ *adjacency matrix* I_a, with rows indexed by complexes and columns indexed by reactions. The entries of I_a are

$$\begin{aligned} -1 &\text{ in row } i \text{ and column } i{\to}j \\ 1 &\text{ in row } j \text{ and column } i{\to}j \\ 0 &\text{ otherwise.} \end{aligned}$$

The reaction rates $\kappa = (\kappa_{ij}) \in \mathbb{R}^r_{>0}$ give the $r \times m$ *weight matrix* I_κ, with rows indexed by reactions and columns indexed by complexes. The entries of I_κ are

$$\begin{aligned} \kappa_{ij} &\text{ in row } i{\to}j \text{ and column } i \\ 0 &\text{ otherwise.} \end{aligned}$$

Thus the ith column of I_κ records the reaction rates of all reactions that *start* from complex i.

The key property of these two matrices is that their product is A_κ, i.e.,

$$A_\kappa = I_a I_\kappa.$$

This enables us to relate the two representations of the ODE system:

$$\begin{aligned}\frac{dx}{dt} &= Y A_\kappa \Phi(x) = Y(I_a I_\kappa)\Phi(x)\\ &= (YI_a)(I_\kappa \Phi(x))\\ &= \begin{bmatrix} \cdots & \overset{i\to j}{\overbrace{y_j - y_i}} & \cdots \end{bmatrix} \begin{bmatrix} \vdots \\ \kappa_{ij} x^{y_i} \\ \vdots \end{bmatrix} \begin{matrix} \\ i\to j \\ \\ \end{matrix}\\ &= \sum_{i\to j} \kappa_{ij} x^{y_i}(y_j - y_i).\end{aligned}$$

Notice also that the columns of YI_a span the stoichiometric subspace S.

Example 5.37. For the Edelstein example

$$\mathrm{A} \underset{\kappa_{21}}{\overset{\kappa_{12}}{\rightleftarrows}} 2\,\mathrm{A} \quad \mathrm{A}+\mathrm{B} \underset{\kappa_{43}}{\overset{\kappa_{34}}{\rightleftarrows}} \mathrm{C} \underset{\kappa_{54}}{\overset{\kappa_{45}}{\rightleftarrows}} \mathrm{B},$$

one easily computes that

$$A_\kappa = \begin{bmatrix} -\kappa_{12} & \kappa_{21} & 0 & 0 & 0 \\ \kappa_{12} & -\kappa_{21} & 0 & 0 & 0 \\ 0 & 0 & -\kappa_{34} & \kappa_{43} & 0 \\ 0 & 0 & \kappa_{34} & -\kappa_{43}-\kappa_{45} & \kappa_{54} \\ 0 & 0 & 0 & \kappa_{45} & -\kappa_{54} \end{bmatrix} =$$

$$\underbrace{\begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}}_{I_a} \underbrace{\begin{bmatrix} \kappa_{12} & 0 & 0 & 0 & 0 \\ 0 & \kappa_{21} & 0 & 0 & 0 \\ 0 & 0 & \kappa_{34} & 0 & 0 \\ 0 & 0 & 0 & \kappa_{43} & 0 \\ 0 & 0 & 0 & \kappa_{45} & 0 \\ 0 & 0 & 0 & 0 & \kappa_{54} \end{bmatrix}}_{I_\kappa}.$$

The fourth column of I_κ shows that two reactions start from C. This will have an interesting consequence. ◁▷

Three Cones. Polyhedral cones have been part of CRNT for a long time. A basic reference is the 1980 paper of Clarke [**79**]. Three cones are especially relevant.

The first is the orthant $\mathbb{R}^n_{\geq 0}$. A positive steady state x_* lies in the interior $\mathbb{R}^n_{>0}$ of this cone. The second cone is more interesting and comes from the Matrix Tree Theorem. When G is weakly reversible, Theorem 5.27 gives an explicit basis $v_1, \ldots, v_\ell$ of $\ker(A_\kappa) \subseteq \mathbb{R}^m$ that is constructed from the partition of $K_1, \ldots, K_m$ induced by the ℓ connected components of G.

Corollary 5.38. *When G is weakly reversible, the basis elements $v_1, \ldots, v_\ell$ of $\ker(A_\kappa) \subseteq \mathbb{R}^m$ are minimal generators of the cone $\ker(A_\kappa) \cap \mathbb{R}^m_{\geq 0}$.*

Proof. For a connected component G_ν of G, $v_\nu = \sum_{i \text{ vertex of } G_\nu} K_i e_i$. The description of K_i given in (5.26) implies $K_i \geq 0$, so that $v_\nu \in \ker(A_\kappa) \cap \mathbb{R}^m_{\geq 0}$. Since the v_ν involve disjoints sets of coordinates and span $\ker(A_\kappa)$, it follows without difficulty that they are minimal generators of the cone. □

The third cone is $\ker(YI_a)\cap\mathbb{R}^r_{\geq 0}$, studied by Clarke in 1980 [**79**]. Here, $\mathbb{R}^r$ has coordinates $z=(z_{ij})_{i\to j}$, indexed by the r reactions of the network, i.e., the edges of G. These are the *reaction coordinates*, sometimes written $z=(z_1,\dots,z_r)$. In [**79**], Clarke uses $\ker(YI_a)\cap\mathbb{R}^r_{\geq 0}$ to analyze the stability of steady states.

To see why the cone $\ker(YI_a)\cap\mathbb{R}^r_{\geq 0}$ is relevant, note that $x_*\in\mathbb{R}^n_{\geq 0}$ implies $I_\kappa\Phi(x_*)\in\mathbb{R}^r_{\geq 0}$. Then for $x_*\in\mathbb{R}^n_{\geq 0}$, we have

$$\begin{aligned} x_* \text{ is a steady state of } (G,\kappa,Y) &\iff 0=YA_\kappa\Phi(x_*)\\ &\iff 0=(YI_a)(I_\kappa\Phi(x_*))\\ &\iff I_\kappa\Phi(x_*)\in\ker(YI_a)\cap\mathbb{R}^r_{\geq 0}\\ &\iff \mathbb{R}_{\geq 0}I_\kappa\Phi(x_*)\subseteq\ker(YI_a)\cap\mathbb{R}^r_{\geq 0}, \end{aligned} \tag{5.37}$$

where the last equivalence follows because $\ker(YI_a)\cap\mathbb{R}^r_{\geq 0}$ is a cone.

Deformed Toric Varieties. The next task is to characterize the elements of $\ker(YI_a)\cap\mathbb{R}^r_{\geq 0}$ that correspond to steady states of (G,κ,Y) for fixed reaction rates $\kappa=(\kappa_{ij})_{i\to j}\in\mathbb{R}^r_{>0}$. We start with the variety

$$V_\kappa = \text{Zariski closure of the image in } \mathbb{C}^r \text{ of } I_\kappa\Phi(x)=(\dots,\underbrace{\kappa_{ij}x^{y_i}}_{i\to j},\dots).$$

We call V_κ a *deformed toric variety* since it is parametrized by monomials multiplied by constants. As we vary κ, we get a family of varieties $V_\kappa\subseteq\mathbb{C}^r$,

The last equivalence of (5.37) suggests using the ray generated by $I_\kappa\Phi(x)$. We capture this idea using a second deformed toric variety:

$$\check{V}_\kappa = \text{Zariski closure of the image in } \mathbb{C}^r \text{ of } x_0I_\kappa\Phi(x)=(\dots,\underbrace{\kappa_{ij}x_0x^{y_i}}_{i\to j},\dots).$$

To find equations for V_κ and $\check{V}_\kappa$, we recast the matrix $Y=[y_1\cdots y_m]\in\mathbb{Z}^{n\times m}_{\geq 0}$ in terms of the reaction coordinates. We do this in two ways:

$$Y_L=\begin{bmatrix} \cdots & \overset{i\to j}{y_i} & \cdots \end{bmatrix}\in\mathbb{Z}^{n\times r}_{\geq 0} \quad\text{and}\quad \check{Y}_L=\begin{bmatrix} \cdots & \overset{i\to j}{1} & \cdots \\ \cdots & y_i & \cdots \end{bmatrix}\in\mathbb{Z}^{(1+n)\times r}_{\geq 0},$$

where some y_i's are omitted (if no reactions start at i) and others appear multiple times (if more than one reaction starts from i).

Using these matrices, it is now easy to find ideals that define the above deformed toric varieties. Recall that any $a\in\mathbb{Z}^r$ can be written uniquely as $a=a^+-a^-$, where $a^+,a^-\in\mathbb{Z}_{\geq 0}$ have disjoint support.

PROPOSITION 5.39. *Let $V_\kappa,\check{V}_\kappa,Y_L,\check{Y}_L$ be as above.*

(1) *If $g_1,\dots,g_t$ is a basis of $\ker(Y_L:\mathbb{Z}^r\to\mathbb{Z}^n)\subseteq\mathbb{Z}^r$, then $V_\kappa=\mathbf{V}(I_{Y_L,\kappa})$ for*

$$I_{Y_L,\kappa}=\left\langle \left(\frac{z}{\kappa}\right)^{g_i^+}-\left(\frac{z}{\kappa}\right)^{g_i^-} \;\middle|\; 1\leq i\leq t\right\rangle : (z_1\cdots z_r)^\infty.$$

(2) *If $h_1,\dots,h_{\check{t}}$ is a basis of $\ker(\check{Y}_L:\mathbb{Z}^r\to\mathbb{Z}^{1+n})\subseteq\mathbb{Z}^r$, then $\check{V}_\kappa=\mathbf{V}(I_{\check{Y}_L,\kappa})$ for*

$$I_{\check{Y}_L,\kappa}=\left\langle \left(\frac{z}{\kappa}\right)^{h_i^+}-\left(\frac{z}{\kappa}\right)^{h_i^-} \;\middle|\; 1\leq i\leq \check{t}\right\rangle : (z_1\cdots z_r)^\infty.$$

PROOF. (1) When $\kappa=\mathbf{1}=(1,\dots,1)$, $V_{\mathbf{1}}$ is the toric variety of Y_L, so that by Exercise 1.1.3 of [**106**], $I_{Y_L,\mathbf{1}}=\langle z^{g_i^+}-z^{g_i^-}\mid 1\leq i\leq t\rangle:(z_1\cdots z_r)^\infty$ is the toric ideal

that defines $\check{V}_{\mathbf{1}}$, The coordinate change given by $z_{i\to j} \mapsto \kappa_{i\to j} z_{i\to j}$ takes $V_{\mathbf{1}}$ to V_κ, and (1) follows easily. The proof of (2) is similar. □

One reason why reaction coordinates are so nice is that changing the reaction rates κ is done by simply rescaling coordinates. This is what makes the proof of Propostion 5.39 so easy.

Let's see what Proposition 5.39 says about the Edelstein example.

EXAMPLE 5.40. $\mathrm{A} \underset{\kappa_{21}}{\overset{\kappa_{12}}{\rightleftarrows}} 2\,\mathrm{A}$, $\mathrm{A}+\mathrm{B} \underset{\kappa_{43}}{\overset{\kappa_{34}}{\rightleftarrows}} \mathrm{C} \underset{\kappa_{54}}{\overset{\kappa_{45}}{\rightleftarrows}} \mathrm{B}$ has reaction coordinates

$$(z_1,\dots,z_6) = (z_{1\to 2}, z_{2\to 1}, z_{3\to 4}, z_{4\to 3}, z_{4\to 5}, z_{5\to 4}),$$

The matrix Y from (5.33) gives matrices

$$Y_L = \begin{bmatrix} 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix} \quad \text{and} \quad \check{Y}_L = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}.$$

Integer bases of the kernels are

$$\begin{aligned} Y_L&\colon \{(0,0,0,-1,1,0),(-1,1,-1,0,0,1),(-2,1,0,0,0,0)\} \\ \check{Y}_L&\colon \{(0,0,0,-1,1,0),(-1,1,-1,0,0,1)\}. \end{aligned}$$

By Proposition 5.39, the second kernel gives the deformed toric ideal

$$\begin{aligned} I_{\check{Y}_L,\kappa} &= \left\langle \frac{z_5}{\kappa_{45}} - \frac{z_4}{\kappa_{43}}, \frac{z_2}{\kappa_{21}}\cdot\frac{z_6}{\kappa_{54}} - \frac{z_1}{\kappa_{12}}\cdot\frac{z_3}{\kappa_{34}} \right\rangle : (z_1\cdots z_6)^\infty \\ &= \langle \kappa_{43}z_5 - \kappa_{45}z_4, \kappa_{12}\kappa_{34}z_2z_6 - \kappa_{21}\kappa_{54}z_1z_3 \rangle \end{aligned}$$

which defines the deformed toric variety $\check{V}_\kappa$. The story for $I_{Y_L,\kappa}$ is similar. □

The Cone and the Deformed Toric Variety. The cone $\ker(YI_a)\cap\mathbb{R}^r_{\geq 0}$ interacts with the deformed toric variety $\check{V}_\kappa$ in two ways. Here is the first interaction, which is a version of [**164**, Lemma 3.1] (see also the comments following the lemma).

THEOREM 5.41. *Suppose $z_* \in \ker(YI_a)\cap\mathbb{R}^r_{>0}$.*

(1) *For any $\kappa > 0$, (G,κ,Y) has a positive steady state x_* with $I_\kappa\Phi(x_*) \in \mathbb{R}_{>0}z_*$ if and only if $z_* \in \check{V}_\kappa$.*
(2) *There exists $\kappa > 0$ such that $z_* \in \check{V}_\kappa$.*

PROOF. One direction of (1) is obvious, since $I_\kappa\Phi(x_*) \in \mathbb{R}_{>0}z_*$ implies there is $\lambda > 0$ such that $z_* = \lambda I_\kappa\Phi(x_*)$. Then $z_* \in \check{V}_\kappa$ by the definition of $\check{V}_\kappa$.

For the other direction, first observe that when $\kappa = \mathbf{1}$, $\check{Y}_L$ induces a monomial map $\mathbb{R}_{>0}\times\mathbb{R}^n_{>0} \to \check{V}_{\mathbf{1}}\cap\mathbb{R}^r_{>0}$ which is easily seen to be surjective by adapting the methods used in the proof of Theorem 5.34.

Given $z_* \in \check{V}_\kappa\cap\ker(YI_a)\cap\mathbb{R}^r_{>0}$, we have $z_*/\kappa \in \check{V}_{\mathbf{1}}\cap\mathbb{R}^r_{>0}$, which by the above paragraph implies the existence of $(\lambda, x_*) \in \mathbb{R}_{>0}\times\mathbb{R}^r_{>0}$ such that

$$z_*/\kappa = (\dots, \underbrace{\lambda x_*^{y_i}}_{i\to j}, \dots).$$

This implies $z_* = \lambda I_\kappa\Phi(x_*)$. Since $z_* \in \ker(YI_a)\cap\mathbb{R}^r_{>0}$, (5.37) implies that x_* is a steady state of (G,κ,Y).

For (2), note that $z_* = x_0 I_\kappa\Phi(x)$ when $\kappa = z_*$, $x_0 = 1$, and $x = (1,\dots,1)$. □

The second interaction between $\check{V}_\kappa$ and $\ker(YI_a) \cap \mathbb{R}^r_{\geq 0}$ involves the minimal integral generators $\{E_i\}$ of this strictly convex rational polyhedral cone. Each E_i lies in the kernel of YI_a, which implies that the E_i come in two flavors:

- $I_a E_i = 0$. It is well known that $\ker(E_i)$ is generated by *positive circuits*, which are directed cycles in G. Being a minimal generator guarantees that E_i comes from a minimal positive circuit.
- $I_a E_i \neq 0$. We call E_i a *stoichiometric generator* in this case.

As explained in the final paragraph of the introduction to [**163**], stoichiometric generators can induce multistationarity.

EXAMPLE 5.42. In the Edelstein example, the integral minimal generators of $\ker(YI_a) \cap \mathbb{R}^6_{\geq 0}$ are

$$E_1 = (1,1,0,0,0,0),\ E_2 = (0,0,1,1,0,0),\ E_3 = (0,0,0,0,1,1),$$
$$E_4 = (1,0,1,0,1,0),\ E_5 = (0,1,0,1,0,1).$$

The first three generators come from the obvious two-cycles in the network shown in Example 5.40. The two other satisfy

$$I_a E_4 = (-1,1,-1,0,1) = -I_a E_5.$$

and hence are stoichiometric generators. The vector $I_a E_4$ lies in the kernel of Y and gives the vector $(-1,1,-1,0,0,1)$ the kernel of $\check{Y}_L$. In Example 5.40, we saw that this vector was responsible for the polynomial $\kappa_{12}\kappa_{34}z_2z_6 - \kappa_{21}\kappa_{54}z_1z_3$ in the deformed toric ideal $I_{\check{V}_L,\kappa}$. ◁▷

The relation between stoichiometric generators and the deformed toric ideal $I_{\check{V}_L,\kappa}$ is explained in [**163**, Lemma 4.5] and the comments following the lemma.

Toric Degenerations. We conclude our discussion of the Edelstein example in [**163**] by examining the toric degenerations of its deformed toric ideal.

EXAMPLE 5.43. For $\mathrm{A} \underset{\kappa_{21}}{\overset{\kappa_{12}}{\rightleftarrows}} 2\,\mathrm{A}$, $\mathrm{A}+\mathrm{B} \underset{\kappa_{43}}{\overset{\kappa_{34}}{\rightleftarrows}} \mathrm{C} \underset{\kappa_{54}}{\overset{\kappa_{45}}{\rightleftarrows}} \mathrm{B}$, we saw in Example 5.40 that the homogeneous deformed toric ideal is

$$I_{\check{Y}_L,\kappa} = \langle \kappa_{43}z_5 - \kappa_{45}z_4, \kappa_{12}\kappa_{34}z_2z_6 - \kappa_{21}\kappa_{54}z_1z_3 \rangle.$$

This ideal I has four initial ideals corresponding to *regular triangulations* of the convex hull of the columns of $\check{Y}_L$. A basic reference is [**333**, Chapter 8].

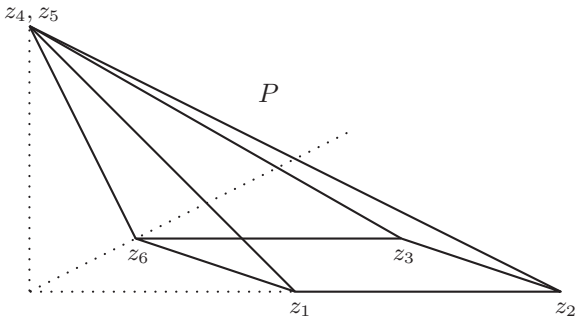
If we write $Y = [y_1\ y_2\ y_3\ y_4\ y_5]$, then $\check{Y}_L$ has six columns

$$\begin{bmatrix}1\\y_1\end{bmatrix}, \begin{bmatrix}1\\y_2\end{bmatrix}, \begin{bmatrix}1\\y_3\end{bmatrix}, \begin{bmatrix}1\\y_4\end{bmatrix}, \begin{bmatrix}1\\y_4\end{bmatrix}, \begin{bmatrix}1\\y_5\end{bmatrix}$$

associated to the reaction variables $z_1, z_2, z_3, z_4, z_5, z_6$. The column containing y_4 gets duplicated since two reactions start from complex 4.

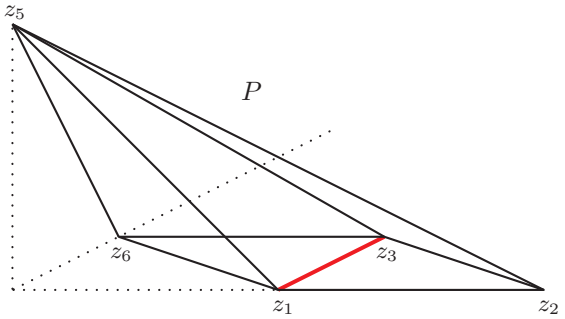
When taking the convex hull, we can ignore the 1 in each column and work with $P = \mathrm{Conv}(y_1, y_2, y_3, y_4, y_4, y_5)$. Labeling each column with the corresponding

reaction coordinate gives the following picture:



We get a regular triangulation by lifting the six vectors into the fourth dimension and projecting the lower convex hull into $\mathbb{R}^3$. Since $y_4 = (0, 0, 1)$ gets lifted twice, the lower convex hull picks z_4 or z_5, depending one which lifting of y_4 is lower. Here is one of the regular trianglulations:

(5.38)



The new edge, shown in red, indicates that the lift of z_2z_6 lies above the lift of z_1z_3. This selects z_2z_6 as the leading term of $\kappa_{12}\kappa_{34}z_2z_6 - \kappa_{21}\kappa_{54}z_1z_3$.

As explained in [**164**, Theorem 4.18], each 3-simplex of the regular triangulation gives a subsystem of the original system. If we label each simplex with the reaction variables of its vertices, we get the subsystems

$$z_1z_2z_3z_5: \quad \mathrm{A} \underset{\kappa_{21}}{\overset{\kappa_{12}}{\rightleftharpoons}} 2\,\mathrm{A} \quad \mathrm{A}+\mathrm{B} \xrightarrow{\kappa_{34}} \mathrm{C} \xrightarrow{\kappa_{45}} \mathrm{B}$$

$$z_1z_3z_5z_6: \quad \mathrm{A} \xrightarrow{\kappa_{12}} 2\,\mathrm{A} \quad \mathrm{A}+\mathrm{B} \xrightarrow{\kappa_{34}} \mathrm{C} \underset{\kappa_{54}}{\overset{\kappa_{45}}{\rightleftharpoons}} \mathrm{B}$$

Let's explore the dynamics of the top subsystem using the reaction rates from (5.31). This gives $\mathrm{A} \underset{1}{\overset{8.5}{\rightleftharpoons}} 2\,\mathrm{A}$, $\mathrm{A}+\mathrm{B} \xrightarrow{1} \mathrm{C} \xrightarrow{1} \mathrm{B}$. To find the steady states, we set $\kappa_{12} = 8.5$, $\kappa_{21} = \kappa_{34} = \kappa_{45} = 1$ and $\kappa_{43} = \kappa_{54} = 0$ in (5.34) to obtain

$$\begin{aligned}
\frac{dx_1}{dt} &= 8.5x_1 - x_1^2 - x_1x_2 \\
\frac{dx_2}{dt} &= -x_1x_2 + x_3 \\
\frac{dx_3}{dt} &= x_1x_2 - x_3.
\end{aligned}$$

When we focus on the stoichiometric compatibility class $x_2 + x_3 = c$ for $c > 0$, we get the system of equations

$$\begin{aligned} 0 &= 8.5x_1 - x_1^2 - x_1x_2 \\ 0 &= -x_1x_2 + x_3 \\ c &= x_2 + x_3. \end{aligned}$$

Elimination theory implies that x_1 satisfies the cubic equation

$$x_1^3 - 7.5x_1^2 + (c - 8.5)x_1 = 0.$$

When $c = 22.5$, it follows that the steady states satisfying $x_2 + x_3 = 22.5$ are

$$(x_1, x_2, x_3) = (0, 22.5, 0) \text{ or } (3.5, 5, 17.5) \text{ or } (4, 4.5, 18).$$

In the first steady state, some concentrations are zero. The other two are positive, which shows that the subsystem selected by the regular triangulation (5.38) exhibits multistationarity. The major result of [**163**] is Theorem 4.18, which relates the behavior of such subsystems to the dynamics of the original system. ◁▷

The idea that the dynamical behavior of a chemical reaction network can be influenced by what happens on certain subsystems goes back to Clarke [**79**] in 1980. This is why he was so interested in cone $\ker(YI_a) \cap \mathbb{R}^r_{\geq 0}$. Gatermann made the link to toric varieties and Gröbner bases in her pioneering work. These days, subsystems of reaction networks are studied using toric varieties, toric ideals, Gröbner bases, and matroids. The papers [**5**, **301**, **321**] and the references therein give a hint of the range and depth of what people are thinking about.

Something Unexpected. We finish this section by observing that the differential equation (5.11) from Section 5.1 can be written as

$$\frac{dx}{dt} = \sum_{i \to j} \underbrace{\kappa_{ij}}_{\text{weights}} \underbrace{x^{y_i}}_{\substack{\text{blending} \\ \text{functions}}} \underbrace{(y_j - y_i)}_{\substack{\text{control} \\ \text{points}}}. \tag{5.39}$$

The language of *blending functions* and *control points* is from geometric modeling, as studied in Section 3.1.

In 2005, Craciun and Feinberg [**111**] related the potential for multistationarity to the injectivity of the map

$$x \in \mathbb{R}^n_{>0} \longmapsto \sum_{i \to j} \kappa_{ij} x^{y_i} (y_j - y_i)$$

for fixed reaction rates $\kappa > 0$. They write this map as $c \mapsto p(c, k)$ and note that

> *If a reaction network has the capacity for multiple positive equilibria, then there exists some choice of positive vector k_0 such that the function $c \mapsto p(c, k_0)$ is not injective.*

Feinberg is one of the creators of CRNT [**147**–**149**], and similarly, Craciun is one of the creators of toric dynamical systems [**113**]. So their collaboration makes sense. But Craicun also wrote the paper *Some geometrical aspects of control points for toric patches* with García-Puente and Sottile [**115**]. This paper uses (5.39) to apply the injectivity theorems of Craciun and Feinberg [**111**, **112**] to geometric modeling. The paper [**115**] also discusses degenerations of toric patches that were studied intensely by García-Puente, Sottile and Zhu in [**160**]. The splendid images presented in Example 3.6 are based on [**160**].

In more recent work [**109**, **110**], Craciun introduced the powerful idea of a *toric differential inclusion* to study the Global Attractor Conjecture from Section 5.1.

There is clearly a lot of interesting mathematics going on here!

Final Comments. In this section, we introduced toric dynamical systems, learned about the work of Karin Gatermann, and spent time on a classic example due to Edelstein. Although our treatment has been far from complete, we hope to have convinced you that this subject is worth learning.

The next section will explore some of the fascinating phenomena that arise in the study of biochemical reaction networks.

5.3. Algebraic Methods for the Study of Biochemical Reaction Networks

by Alicia Dickenstein

In this section we will concentrate on biochemical reaction networks of interest in systems biology, in particular enzymatic networks, consisting of different types of multisite phosphorylation networks. One source of inspiration for our study with algebro-geometric tools is the following quote from the abstract of the paper [**300**]:

> Multisite phosphorylation cycles are ubiquitous in cell regulation systems and are studied at multiple levels of complexity, from molecules to organisms, with the ultimate goal of establishing predictive understanding of the effects of genetic and pharmacological perturbations of protein phosphorylation in vivo. Achieving this goal is essentially impossible without mathematical models, which provide a systematic framework for exploring dynamic interactions of multiple network components.

We will mainly concentrate on recent advances on the determination of multistationarity for these networks, whose dynamics are usually modeled with mass-action kinetics. For many classes of chemical networks, as the complex balanced networks discussed in Sections 5.1 and 5.2, monostationarity is an important property. Instead, for biochemical reaction networks, that is, chemical reaction networks modeling pathways in systems biology, multistationarity is a general feature and it is important because it is intepreted as a way for the cell *to take different decisions*. Indeed, differential systems with mass-action kinetics are deterministic. But the occurrence of multiple stable steady states in the same stoichiometric compatibility class[1] implies that trajectories starting from different initial conditions with the *same* conserved quantities can converge to steady states with different properties.

We will end the chapter with some open questions. We refer the reader to the surveys [**86**, **88**, **126**, **127**] for further details and references.

Multistationarity. We have seen the definition of multistationarity of a given reaction network in Section 5.1. Our point of view will be the following. We will fix a basis $\mathcal{B}$ of the orthogonal $S^{\perp}$ of the stoichiometric subspace S, and we call *total amounts* the linear conservation constants associated to a trajectory x, that is, the values c_v of the conserved quantities $v \cdot x$, for any $v \in \mathcal{B}$. Given a reaction network, we will consider both the vector κ of reaction rate constants and the vector $c = (c_v)_{v \in \mathcal{B}}$ of total amounts as *parameters*. We will thus look at the family of systems of differential equations of the form (5.11) (parametrized by κ)

[1] Such steady states are said to be *stoichiometrically compatible*

and stoichiometric compatibility classes $\{v \cdot x = c_v, v \in \mathcal{B}\} \cap \mathbb{R}^n_{\geq 0}$. Note that $S_c := \{v \cdot x = c_v, v \in \mathcal{B}\} = x(0) + S$, for any $x(0)$ satisfying $v \cdot x(0) = c_v$, for all $v \in \mathcal{B}$, as described in (5.16). Recall that we only allow the rate constants κ to take positive values.

DEFINITION 5.44. We say that a given reaction network has the *capacity of multistationarity* if there exists a choice of parameters κ, c for which the system with rate constants κ has more than one *positive* steady state in the stoichiometric compatibility class $S_c \cap \mathbb{R}^n_{>0}$. Any such choice of κ, c are called *multistationarity parameters*. When there is at most a single positive steady state in each class, the reaction network is called *monostationary*.

Main Questions About Multistationarity. Given a reaction network, these are the main questions about multistationarity:

(1) Does the network have the capacity of multistationarity?
(2) If so,
   (a) detect multistationarity parameters. The answer to this query could be just one choice of parameters, an open set in the space of parameters (called a *multistationarity region*) or hopefully, all the parameters for which multistationarity occurs.
   (b) find the maximum possible number of positive steady states in the same stoichiometric compatibility class, or at least, give a refined lower bound for this maximum number. Detect multistationarity parameters with maximum/many stoichiometrically compatible positive steady states.

As we remarked in (5.17), the steady state variety V^{ss} is an algebraic variety, as well as the translates S_c. So the previous questions are indeed questions about the number of positive solutions of multivariate sparse systems of real polynomials. This is a question of quantifier elimination in real algebraic geometry, which is in principle effectively computable for a given system, but with a high complexity (in particular for parametric systems). Giving sharp bounds on the number of positive zeros for parametric polynomial systems from their structure is a very difficult theoretical problem. One obvious upper bound is the Bernstein-Kouchnirenko-Khovanskii bound for the number of complex solutions with nonzero coordinates (see for instance p. 206 in [**165**]), which is usually much bigger than the number of positive solutions. The main result in the area is due to Khovanskii [**233**]. He proved that given a sparse system of n polynomials in n variables with real coefficients and exponents in a finite set A, there is a bound for the number of nondegenerate positive real roots in terms of the cardinality of A. This bound is very far from being sharp. There are some sharper particular results (see e.g. [**31**]), but the general question of finding a tight bound is out of reach for the moment.

We now present some important basic biological mechanisms.

EXAMPLE 5.45. **Phosphorylation Cycles.** Multisite protein phosphorylation/dephosphorylation plays a prominent role in intracellular processes like signal transduction and cell-cycle control. The distributive sequential n-phosphorylation cycle has been extensively studied. It describes the phosphorylation of a protein with n sites where a phosphate group can be added or released (we denote these species by $S_0, S_1, \ldots, S_n$, where S_i means that it has i sites with a phosphate group) by a pair of enzymes (a kinase E and a phosphatase F) in a sequential

and distributive mechanism. This means that there are intermediate species at each phospho/dephosphorylation step, where only one phosphate group is added or released. The reaction network for the case $n = 2$ is as follows:

$$S_0 + E \underset{k_{\mathrm{off}_0}}{\overset{k_{\mathrm{on}_0}}{\rightleftarrows}} ES_0 \xrightarrow{k_{\mathrm{cat}_0}} S_1 + E \underset{k_{\mathrm{off}_1}}{\overset{k_{\mathrm{on}_1}}{\rightleftarrows}} ES_1 \xrightarrow{k_{\mathrm{cat}_1}} S_2 + E$$

$$S_2 + F \underset{l_{\mathrm{off}_1}}{\overset{l_{\mathrm{on}_1}}{\rightleftarrows}} FS_2 \xrightarrow{l_{\mathrm{cat}_1}} S_1 + F \underset{l_{\mathrm{off}_0}}{\overset{l_{\mathrm{on}_0}}{\rightleftarrows}} FS_1 \xrightarrow{l_{\mathrm{cat}_0}} S_0 + F$$

We denote the concentrations of the nine chemical species S_0, S_1, S_2, E, F, and ES_0, ES_1, FS_2, FS_1, by x_1, x_2, x_3, x_4, x_5, and y_1, y_2, y_3, y_4, respectively. The associated mass-action kinetics dynamical system describing the evolution in time of the concentrations equals:

$$\begin{aligned}
\frac{dx_1}{dt} =& -k_{\mathrm{on}_0} x_1 x_4 + k_{\mathrm{off}_0} y_1 + l_{\mathrm{cat}_0} y_4 \\
\frac{dx_2}{dt} =& -k_{\mathrm{on}_1} x_2 x_4 + k_{\mathrm{cat}_0} y_1 + k_{\mathrm{off}_1} y_2 \\
& - l_{\mathrm{on}_0} x_2 x_5 + l_{\mathrm{cat}_1} y_3 + l_{\mathrm{off}_0} y_4 \\
\frac{dx_3}{dt} =& k_{\mathrm{cat}_1} y_2 - l_{\mathrm{on}_1} x_3 x_5 + l_{\mathrm{off}_1} y_3 \\
\frac{dy_1}{dt} =& k_{\mathrm{on}_0} x_1 x_4 - (k_{\mathrm{off}_0} + k_{\mathrm{cat}_0}) y_1 \\
\frac{dy_2}{dt} =& k_{\mathrm{on}_1} x_2 x_4 - (k_{\mathrm{off}_1} + k_{\mathrm{cat}_1}) y_2
\end{aligned}
\qquad
\begin{aligned}
\frac{dx_4}{dt} =& -k_{\mathrm{on}_0} x_1 x_4 - k_{\mathrm{on}_1} x_2 x_4 + (k_{\mathrm{off}_0} + k_{\mathrm{cat}_0}) y_1 \\
& + (k_{\mathrm{off}_1} + k_{\mathrm{cat}_1}) y_2 \\
\frac{dx_5}{dt} =& -l_{\mathrm{on}_0} x_2 x_5 - l_{\mathrm{on}_1} x_3 x_5 + (l_{\mathrm{off}_1} + l_{\mathrm{cat}_1}) y_3 \\
& + (l_{\mathrm{off}_0} + l_{\mathrm{cat}_0}) y_4 \\
\frac{dy_3}{dt} =& l_{\mathrm{on}_1} x_3 x_5 - (l_{\mathrm{off}_1} + l_{\mathrm{cat}_1}) y_3 \\
\frac{dy_4}{dt} =& l_{\mathrm{on}_0} x_2 x_5 - (l_{\mathrm{off}_0} + l_{\mathrm{cat}_0}) y_4 .
\end{aligned}$$

The space of linear conservation relations has dimension 3. The basis which is usually taken gives rise to the following linear invariants of the trajectories:

$$\begin{aligned}
x_1 + x_2 + x_3 + y_1 + y_2 + y_3 + y_4 =& S_{\mathrm{tot}} \\
x_4 + y_1 + y_2 =& E_{\mathrm{tot}} \\
x_5 + y_3 + y_4 =& F_{\mathrm{tot}},
\end{aligned}$$

where $S_{\mathrm{tot}}, E_{\mathrm{tot}}, F_{\mathrm{tot}}$ are positive real numbers for any choice of initial condition in the positive orthant. Note that there are $12 + 3 = 15$ parameters. For any $n \in \mathbb{N}$, the system is similar, with $3n + 3$ species (so $3n + 3$ variables, corresponding to the concentrations of $n + 1$ substrates $S_0, \ldots, S_n$, 2 enzymes E, F and $2n$ intermediate species), and $6n + 3$ parameters ($6n$ reactions and so, $6n$ reaction rate constants, plus 3 independent conservation relations and so, 3 total amounts). Notably, there are many general results even with this number of variables and parameters. We will outline the methods to get these results below.

When $n \geq 2$, the distributive sequential n-phosphorylation cycle has the capacity of multistationarity. This is shown in the nice article [**357**] by Wang and Sontag, where it is proved that there cannot be more than $2n - 1$ positive steady states, while there are parameters for which the system has $n + 1$ positive steady states in case n is even and n in case n is odd (that is, $2\lfloor n/2 \rfloor + 1$, where $\lfloor \cdot \rfloor$ denotes integer part). Multistationarity for $n \geq 2$ is also explored in [**209**]. In [**155**] it is shown that the upper bound $2n - 1$ is attained for certain parameter values in case $n = 3, 4$ (this also holds for $n = 2$ since $2 \cdot 2 - 1 = 2 + 1$). It is an open question whether this upper bound is attained for $n \geq 5$!

In the paper [**30**], parameter regions on all the parameters are given for the occurrence of multistationarity for the distributive sequential n-site phosphorylation systems (with 3 positive steady states), using a polyhedral approach that we briefly describe below. Using these same tools, the paper [**168**] describes regions of multistationarity for any n where the number of stoichiometric compatible positive steady states equals $2\lfloor n/2 \rfloor + 1$. On the other hand, in the article [**150**] a choice of parameters is found for which there are this number of nondegenerate stoichiometrically compatible steady states and moreover, $\lfloor n/2 \rfloor + 1$ are *stable*, with techniques of singular perturbation theory (it had been shown evidence for this unlimited mulstistability in [**346**]). When there are no linear conservation relations, a (nondegenerate) steady state x_* is *stable* if all the the eigenvalues of the Jacobian of $f_1, \ldots, f_n$ at x_* have negative real part. In case the stoichiometric subspace $S \neq 0$ (and the linear forms in $S^\perp$ define all linear relations among $f_1, \ldots, f_n$, a (nondegenerate) steady state is *stable* if the Jacobian at x_* has $n - \dim(S)$ eigenvalues equal to 0 and the rest have negative real part. In this case, trajectories sufficiently close to x_* will converge to it.

In the *processive* n-site phosphorylation cycle instead, the unphosphorylated substrate S_0 binds to the enzyme E and eventually gets n-phosphorylated sites (S_n) via monomolecular intermediate complexes consisting of a single intermediate species $U_1, \ldots, U_n$, and the dephosphorylations reactions are modeled similarly (through intermediate species $V_n, \ldots, V_1$ by the action of the enzyme F):

$$\mathrm{S}_0 + \mathrm{E} \underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} \mathrm{U}_1 \xrightarrow{\kappa_3} \mathrm{U}_2 \xrightarrow{\kappa_4} \cdots \xrightarrow{\kappa_{n+1}} \mathrm{U}_n \xrightarrow{\kappa_{n+2}} \mathrm{S}_n + \mathrm{E}$$

$$\mathrm{S}_n + \mathrm{F} \underset{\ell_2}{\overset{\ell_1}{\rightleftharpoons}} \mathrm{V}_n \xrightarrow{\ell_3} \mathrm{V}_{n-1} \longrightarrow \cdots \xrightarrow{\ell_{n+1}} \mathrm{V}_1 \xrightarrow{\ell_{n+2}} \mathrm{S}_0 + \mathrm{F}.$$

In this case, there also 3 linearly independent conservation relations, $2n+4$ species and $(2n+4)+3 = 2n+7$ parameters. It is proved in [**87**] that the system is monostationary for any value of n and moreover, that the only steady state in each stoichiometric compatibility class is a global attractor for all trajectories starting in the positive orthant. We refer the reader to the survey [**88**] for futher recent progress and directions. ◁▷

EXAMPLE 5.46. **Enzymatic Cascades.** The simplest case of a cascade with the capacity of multistationarity consists of a cascade with two layers and a single phosphorylation/dephosphorylation at each layer, with one phosphatase [**152**]. It corresponds to a labeled directed graph, with 9 variables and 18 parameters, where each phosphorylation follows the same mechanism as in the processive distributive cycle in Example 5.45, with an intermediate species. The ten species are the substrates S_0, S_1 in the first layer, the substrates P_0, P_1 in the second layer, four intermediate complexes, a kinase E and the *same* phosphatase F which dephosphorylates the substrates in both layers. The forward enzyme in the second layer is the phosphorylated substrate S_1 from the first layer.

The cascade is shown on the next page on the left. On the right, we show a typical depiction of the mechanism that hides the reaction rate constants and the intermediate species.

In this case, there are 4 linearly independent conservation relations. In case there are two different phosphatases F_1, F_2 that interact with S_1 and P_1 in the dephosphorylation path, the system has 5 linearly independent conservation relations and is monostationary for any choice of parameters. In [**167**] we described regions

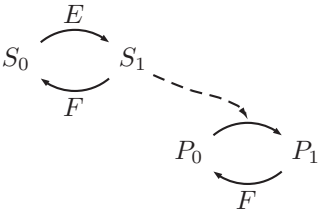
of multistationarity in terms of all the parameters for cascades with n-layers (and so with an increasing unbounded number of linearly independent conservation relations), when there is the same phosphatase F acting in two of these layers. Again, we used for this the polyhedral approach in [**30**].

$$S_0 + E \underset{k^1_{\mathrm{off}_0}}{\overset{k^1_{\mathrm{on}_0}}{\rightleftarrows}} ES_0 \xrightarrow{k^1_{\mathrm{cat}_0}} S_1 + E$$

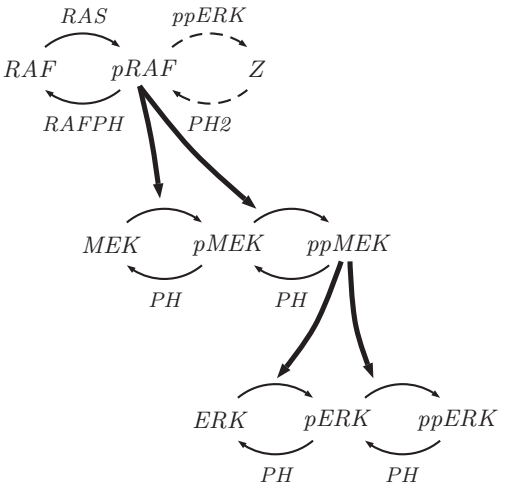
$$S_1 + F \underset{l^1_{\mathrm{off}_0}}{\overset{l^1_{\mathrm{on}_0}}{\rightleftarrows}} FS_1 \xrightarrow{l^1_{\mathrm{cat}_0}} S_0 + F$$

$$P_0 + S_1 \underset{k^2_{\mathrm{off}_0}}{\overset{k^2_{\mathrm{on}_0}}{\rightleftarrows}} S_1P_0 \xrightarrow{k^2_{\mathrm{cat}_0}} P_1 + S_1$$

$$P_1 + F \underset{l^2_{\mathrm{off}_0}}{\overset{l^2_{\mathrm{on}_0}}{\rightleftarrows}} FP_1 \xrightarrow{l^2_{\mathrm{cat}_0}} P_0 + F.$$



Another important enzymatic cascade is the extracellular signal-regulated kinase (ERK) pathway, which leads to activation of the effector molecule ERK, which controls downstream responses in the cell. It has a crucial role in the occurrence and development of diseases or developmental defects [**285**]. The protein RAS is produced inside the cell after a mechanism that started through the membrane, and in turn phosphorylates RAF, which activates MEK and then it allows for the double phosphorylation of ERK.



The dotted arrows indicate a possible retroactivity in the cascade by the activated molecule ppERK. This network admits multistationarity, but it is an open problem to determine which is the maximum possible number of stoichiometrically compatible positive steady states. The number of complex solutions to the equations $f_i(x) = 0$ and the linear equations defining a stoichiometric compatibility class can be computed, but this is certainly a non-sharp upper bound for the number of positive solutions. Even if there are *only* three layers, the number of variables and parameters is too big to use in a direct way any theoretical approach, as for

instance via the computation of discriminants [**165**]. It might be possible to computationally attack this question via an improvement of the RAGlib library [**304**]. ◁▷

EXAMPLE 5.47. **Bifunctional Enzymes and ACR.** We now consider an osmolarity regulation network in bacteria, extracted from [**320**], which is an example of a so-called *two-component* system. There are two proteins, the sensor kinase EnvZ (denoted by X_1) and the response regulator OmpR (denoted by Y_1). EnvZ autophosphorylates (X_3) and catalyzes the *swap* of the phosphate group to the response regulator (the phosphorylated form OmpR-P is denoted by Y_2) via an intermediate species U_1. In this mechanism, EnvZ can also be bound to ATP (X_2) or to ADP (X_0), and in both cases it acts as an enzyme, dephosphorylating Y_2 back to Y_1 via intermediate species U_2 and U_3. The associated directed graph is as follows:

$$\mathrm{X}_0 \underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} \mathrm{X}_1 \underset{\kappa_4}{\overset{\kappa_3}{\rightleftharpoons}} \mathrm{X}_2 \xrightarrow{\kappa_5} \mathrm{X}_3$$

$$\mathrm{X}_3 + \mathrm{Y}_1 \underset{\kappa_7}{\overset{\kappa_6}{\rightleftharpoons}} \mathrm{U}_1 \xrightarrow{\kappa_8} \mathrm{X}_1 + \mathrm{Y}_2$$

$$\mathrm{X}_2 + \mathrm{Y}_2 \underset{\kappa_{10}}{\overset{\kappa_9}{\rightleftharpoons}} \mathrm{U}_2 \xrightarrow{\kappa_{11}} \mathrm{X}_2 + \mathrm{Y}_1$$

$$\mathrm{X}_0 + \mathrm{Y}_2 \underset{\kappa_{13}}{\overset{\kappa_{12}}{\rightleftharpoons}} \mathrm{U}_3 \xrightarrow{\kappa_{14}} \mathrm{X}_0 + \mathrm{Y}_1$$

As usual, we denote the concentrations of the species with small letters. There are two linearly independent conservation relations:

$$x_0 + x_1 + x_2 + x_3 + u_1 + u_2 + u_3 = X_{\mathrm{tot}},\ y_1 + y_2 + u_1 + u_2 + u_3 = Y_{\mathrm{tot}}. \tag{5.40}$$

The steady state variety V^{ss} can then cut out in $\mathbb{R}^9$ by 7 linearly independent polynomials among $f_1, \ldots, f_9$ and for positive rate constants has dimension two. The particular design of this biochemical network keeps for any positive steady state the concentration y_2 constant with value $\alpha(\kappa) > 0$ described in (5.41) below, which depends on the reaction rate constants but is independent of the positive initial concentrations (as long as $Y_{\mathrm{tot}} > \alpha(\kappa)$). It is then said that the system shows *Absolute Concentration Robustness* (ACR) in the concentration y_2 of the phosphorylated response regulator. We show in [**291**] that it is not possible to detect this only via linear combinations of the polynomials $f_1, \ldots, f_9$ with coefficients in the field $\mathbb{R}$. But this becomes evident *for generic choices of the rate constants* via the computation of a Gröbner basis of the ideal I_f spanned by $f_1, \ldots, f_9$ in the polynomial ring $\mathbb{Q}(\kappa_1, \ldots, \kappa_{14})[x_0, x_1, x_2, x_3, y_1, y_2, u_1, u_2, u_3]$ with coefficients in the field of rational functions of the reaction rate constants.

We now explain the elimination of intermediate species developed in [**153**] in this case, which allows us to do the computation more easily. Order the variables $x = (x_0, x_1, x_2, x_3, y_1, y_2, u_1, u_2, u_3)$. Then

$$\begin{array}{rcccl} \frac{du_1}{dt} & = & f_7(x) & = & \kappa_6\, x_3 y_1 - (\kappa_7 + \kappa_8) u_1 \\ \frac{du_2}{dt} & = & f_8(x) & = & \kappa_9\, x_2 y_2 - (\kappa_{10} + \kappa_{11}) u_2 \\ \frac{du_3}{dt} & = & f_9(x) & = & \kappa_{12}\, x_0 y_2 - (\kappa_{13} + \kappa_{14}) u_3. \end{array}$$

It follows that at steady state, we can compute the values of the concentrations u_1, u_2, u_3 from the values of the concentrations x_0, x_2, x_3, y_1, y_2 and the rate constants. Also, define the following rational functions of κ, which are well defined for

any choice of positive values of the reaction rate constants:

$$\tau_1 = \frac{\kappa_6\kappa_8}{\kappa_7+\kappa_8},\ \tau_2 = \frac{\kappa_9\kappa_{11}}{\kappa_{10}+\kappa_{11}},\ \tau_3 = \frac{\kappa_{12}\kappa_{14}}{\kappa_{13}+\kappa_{14}}.$$

Then, if we skip the concentrations of the intermediate species and we order the variables $x' = (x_0, x_1, x_2, x_3, y_1, y_2)$, let $\dfrac{dx'}{dt} = g(x')$ be the mass-action kinetics system associated with the reduced network:

$$\begin{aligned}
&\mathrm{X}_0 \underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} \mathrm{X}_1 \underset{\kappa_4}{\overset{\kappa_3}{\rightleftharpoons}} \mathrm{X}_2 \xrightarrow{\kappa_5} \mathrm{X}_3\\
&\mathrm{X}_3 + \mathrm{Y}_1 \xrightarrow{\tau_1} \mathrm{X}_1 + \mathrm{Y}_2\\
&\mathrm{X}_2 + \mathrm{Y}_2 \xrightarrow{\tau_2} \mathrm{X}_2 + \mathrm{Y}_1\\
&\mathrm{X}_0 + \mathrm{Y}_2 \xrightarrow{\tau_3} \mathrm{X}_0 + \mathrm{Y}_1
\end{aligned}$$

Here, $x \in V^{ss}$ if and only if

$$x = \big(x', \frac{\kappa_6}{\kappa_7+\kappa_8}x_3y_1, \frac{\kappa_9}{\kappa_{10}+\kappa_{11}}x_2y_2, \frac{\kappa_{12}}{\kappa_{13}+\kappa_{14}}x_0y_2\big), \quad \text{with } g(x') = 0.$$

Now, $g = (g_1, \ldots, g_6)$ satisfy two linearly independent linear relations $g_1 + g_2 + g_3 + g_4 = 0$, $g_5 + g_6 = 0$, and thus the ideal I_g is spanned, for instance, by the polynomials:

$$\begin{aligned}
g_1 &= -\kappa_1 x_0 + \kappa_2 x_1\\
g_3 &= \kappa_3 x_1 - (\kappa_4+\kappa_5)x_2\\
g_4 &= \kappa_5 x_2 - \tau_1 x_3 y_1\\
g_5 &= -\tau_1 x_3 y_1 + \tau_2 x_2 y_2 + \tau_3 x_0 y_2.
\end{aligned}$$

If we compute a Gröbner basis of I_g with respect to the purely lexicographic order where (x_2, y_2) are the smaller variables (directly or via the FGLM algorithm), we find the following polynomial

$$h = x_2((\kappa_1\,\tau_2\,\kappa_3 + \tau_3\,\kappa_2\,\kappa_4 + \tau_3\,\kappa_2\,\kappa_5)\,y_2 - \kappa_1\,\kappa_3\,\kappa_5),$$

and we deduce that when $x_2 \neq 0$, the value of y_2 at steady state equals the constant

$$\alpha(\kappa) = \frac{\kappa_1\,\kappa_3\,\kappa_5}{\kappa_1\,\tau_2\,\kappa_3 + \tau_3\,\kappa_2\,\kappa_4 + \tau_3\,\kappa_2\,\kappa_5}, \tag{5.41}$$

which is well defined since all $\kappa_i, \tau_i > 0$ and depends only on the reaction rate constants. Note that when computing a Gröbner basis with a term order that only depends on the variables x' and not on the parameters, we are indeed working in the polynomial ring with coefficients in $\mathbb{Q}(\kappa)$. We invite the reader to check that this is in fact a binomial ideal in the variables x' by checking that the reduced Gröbner basis consists of binomials [**141**], but that the ideal generated by $g_1, \ldots, g_6$ in the polynomial ring $\mathbb{Q}[\kappa_1, \ldots, \kappa_5, \tau_1, \tau_2, \tau_3, x']$ is not binomial and the value of y_2 at a zero of the ideal is not constant when the variables x' are nonzero. It follows that the ideal generated by $f_1, \ldots, f_9$ in $\mathbb{Q}(\kappa)[x]$ is binomial but the ideal it generates in $\mathbb{Q}[\kappa, x]$ is not.

This network also has the following property: it has *relevant boundary steady states*, which is not shared by any of the reaction networks we presented in Examples 5.45 and 5.46. A boundary steady state in a system with n variables is a point $x \in V^{ss}_{\geq 0} \setminus \mathbb{R}^n_{>0}$ (see (5.18)), that is, a steady state with nonnegative coordinates and at least one coordinate equal to 0. A boundary steady state is

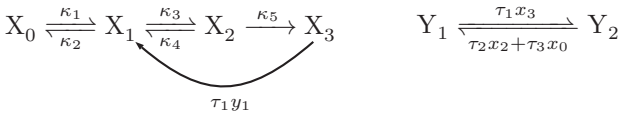
said to be *relevant* if it lies in a stoichiometric compatibility class which intersects $\mathbb{R}^n_{>0}$. Indeed, the stoichiometric compatibility class defined by the affine relations (5.40) intersects the positive orthant if and only if $Y_{\rm tot} > \alpha(\kappa)$. In this case, $x = (0, 0, 0, X_{tot}, 0, Y_{tot}, 0, 0, 0)$ is a relevant boundary steady state. The cardinality of $V^{ss}_{\geq 0}$ is 2, but only one of the steady states lies in the open orthant $\mathbb{R}^{10}_{>0}$ and so the network is monostationary. ◁▷

General Results From Structure. In recent years, several authors developed general results for biochemical reaction networks with algebro-geometric tools from the structure of the network, in particular enzymatic networks as those presented in Examples 5.45 and 5.46, starting with the paper [**345**] by Thomson and Gunawardena. Other articles with this perspective are the papers [**173**, **174**] by Gnacadja, and the article [**153**] by Feliu and Wiuf. Together with Pérez Millán, and building on this previous work, we identified in [**290**] a common structure in many popular biological networks, called a MESSI structure, that describes *Modifications of type Enzyme-Substrate or Swap with Intermediates*. All the examples we presented are instances of MESSI systems.

The main idea is that the set $\mathcal{S}$ of species can be naturally partitioned into a (possibly empty) set $\mathcal{S}^{(0)}$ of *intermediate* species and several (nonempty) sets $\mathcal{S}^{(1)}, \ldots, \mathcal{S}^{(m)}$ of *core* species. Each subset $\mathcal{S}^{(\alpha)}$ of core species has its associated intermediate species $Int(\alpha)$ and the mass-action kinetics system is (affine) linear in the variables in $\mathcal{S}^{(\alpha)} \cup Int(\alpha)$. These linearities inside the nonlinear system are key to proving different properties. Note that the subsets $Int(\alpha)$ are not disjoint, and this interaction is many times responsible for the multistationarity of the system. We refer the reader to [**290**] for the general definition. There could be different partitions of the set of species inducing a MESSI structure on a given network; in fact, there is a poset of them and one could always consider a minimal one.

Given a partition defining a MESSI structure on a biochemical reaction network with underlying directed graph G of reactions, we can associate to it three other directed graphs G_1, G_2, G_E in an algorithmic way. First, G_1 corresponds to the elimination of intermediates, modifying the reaction rate constants as in our Example 5.47, based on the results of [**153**]. In Example 5.47, a (minimal) partition is given by $\mathcal{S}^{(0)} = \{U_1, U_2, U_3\}$, $\mathcal{S}^{(1)} = \{X_0, X_1, X_2, X_3\}$, $\mathcal{S}^{(2)} = \{Y_1, Y_2\}$, and $Int(1) = Int(2) = \mathcal{S}^{(0)}$. The underlying graph of the reduced network in this example is the directed graph G_1. We also refer the reader to the papers [**151**, **303**] for the *positive* elimination of variables in the steady state equations.

The directed graph G_2 has, in the case of minimal partitions, one connected component for each subset $\mathcal{S}^{(\alpha)}, \alpha = 1, \ldots, m$, and in the component corresponding to $\mathcal{S}^{(\alpha)}$, all the species not in this subset of the partition are "hidden" in the labels of the directed graph. Again in the case of Example 5.47, which besides two enzymatic reactions has a swap, the corresponding directed graph G_2 equals:



Formally, the mass-action systems associated with G_1 and G_2 define the same steady state equations (and thus, the same steady state ideal).

Finally, the directed graph G_E has m nodes corresponding to the subsets of core species in the partition and there is an arrow from $\mathcal{S}^{(\alpha)}$ to $\mathcal{S}^{(\beta)}$ if and only if at least one species from the first subset occurs as a label in the component associated to the second subset in G_2 (after removal of loops, if any). The edges in G_E are not labeled. In the case of Example 5.47, the corresponding graph G_E equals

$$\mathcal{S}^{(1)} \rightleftharpoons \mathcal{S}^{(2)}. \tag{5.42}$$

Theorem 3.2 in [**290**] proves that a MESSI system has one (independent) linear conservation relation associated to each of the subsets in the partition of the species set corresponding to core species (as in (5.40)). This implies that all stoichiometric compatibility classes are bounded and all trajectories are defined for any $t \in \mathbb{R}_{\geq 0}$. Moreover, sufficient conditions are given to ensure that there are no other independent linear relations. This gives the bases of linear relations in the networks in Examples 5.45 and 5.46. Theorem 3.15 in [**290**] asserts that when G_2 is weakly reversible and G_E has no directed cycles, then there cannot be relevant boundary steady states and so any trajectory starting in the positive orthant is at a positive distance from the boundary of the orthant (this is called *persistence*). Note that in (5.42), the directed graph G_E does have a cycle and there are relevant boundary steady states, so this result does not apply to this case, but all the enzymatic networks in our first two examples do satisfy the conditions of the theorem. Theorem 4.1 in [**290**] gives sufficient conditions for $V^{ss} \cap \mathbb{R}^n_{>0}$ to have a rational parametrization, in a very explicit and algorithmic way. This is even more explicit for *s-toric* MESSI systems, in which $V^{ss} \cap \mathbb{R}^n_{>0}$ can be cut out by binomials read from the structure of the network (see Definition 4.3 and Theorem 4.8 in [**290**]). Moreover, in this case, it is possible to exactly characterize the capacity for multistationarity and give choices of multistationarity parameters.

Below we state a version of Theorem 5.4 in [**290**]. This result, in terms of the associated *oriented matroids* of the matrix of exponents and the matrix of coefficients of the system, is based on many previous results, starting with those in the paper [**111**] by Craciun and Feinberg, and ending with the comprehensive paper [**277**], including several others [**278**, **291**]. We keep the previous notation; in particular, κ denotes the vector of reaction rate constants of a given network on n species. Given a matrix with n rows (resp. n columns) and a subset $J \subseteq \{1, \ldots, n\}$ we indicate with the subscript J the submatrix obtained considering only the rows (resp. the columns) indicated by the elements of J. We denote $\sigma(J) = \sum_{j \in J} j$.

THEOREM 5.48. *Assume that the steady state variety $V = V^{ss} \cap \mathbb{R}^n_{>0}$ of a chemical reaction network is nonempty and cut out by binomials with coefficients in $\mathbb{Q}(\kappa)$ that are defined for positive values of the rate constants, and exponents arranged as columns of a matrix, from which we choose a submatrix B of maximal rank. Equivalently, $V \neq \emptyset$ is parametrized by monomials with exponents in any dual matrix A of maximal rank such that the columns of B generate its kernel. Let $S^\perp$ denote a matrix whose rows define the dual of the stoichiometric subspace S. If $\operatorname{rank}(S^\perp) = \operatorname{rank}(A) = d$, then the following statements are equivalent:*

(1) *Monostationarity: for any $x(0)$ in $\mathbb{R}^n_{>0}$ and any choice of rate constants κ, there is at most a single steady state in $V \cap (x(0) + S)$.*
(2) *For all subsets $J \subseteq \{1, \ldots, n\}$ of cardinality d, the product $\det(S^\perp_J)\det(A_J)$ either is zero or has the **same sign** as all other nonzero products, and at least one such product is nonzero.*

(3) *Same sign conditions with* $\det(S_J)\det(B_J)$.
(4) *Same sign conditions with* $(-1)^{\sigma(J)}\det(S_J^{\perp})\det(B_{\{1,\dots,n\}\setminus J})$.

The equivalent statements in Theorem 5.48 are also equivalent to the fact that there are no vectors $w \in S$ and $v \in \operatorname{rowspan}(A)$ which have coordinatewise the same sign (see for instance [**279**]). When such vectors do exist, they can be used to construct multistationarity parameters κ (as in [**290**, **291**]). The corresponding total amounts are then computed from the steady states. All the results we mentioned are algorithmic and there will be soon a publicly available open source free implementation by Guillermo Mosse, based on Algorithm 1 in [**290**].

There is another equivalent condition to the four equivalent statements in Theorem 5.48 in terms of the signs of the coefficients of a polynomial with $(0, 1)$ exponents which can be computed as a determinant. We refer to [**277**, **363**] for details about these determinant conditions and to [**302**] for a use of this approach to detect minimal subsets of intermediate species that allow for multistationarity.

Other Approaches to the Question of Multistationarity. We quickly review other approaches to the question of multistationarity.

Using Brouwer Degree Theory: Another interesting approach to decide the capacity for multistationarity without parameter sampling is based on degree theory. We refer the reader to Theorem 1 in [**83**]. In this paper, the authors introduce a procedure to partition the reaction rate space into regions for which the system has a unique or multiple equilibria. Their procedure is applied to several models of gene transcription and cell signalling, for which a complete partitioning of the reaction rate space with respect to multistationarity can be obtained. This approach does not provide conditions on the total amounts. It has been simplified for so called *linearly binomial networks* (which include many s-toric MESSI systems) in [**131**].

Using Polyhedral Methods: This approach is based on the paper [**32**] to get lower bounds for the number of real positive roots of sparse systems of polynomials, by lifting positive real roots from a degeneration to smaller systems that correspond to simplices in a *regular subdivision* of the support of the polynomials. We refer the reader to [**30**] for general results on the use of these ideas to find *regions of multistationarity* in MESSI systems with a rational parametrization, as the enzymatic examples in Examples 5.45 and 5.46. We also refer to Section 3 in [**167**] for a summary and a clarifying example. This approach provides sufficient (but not necessary) open conditions in the space of all parameters (reaction rate constants and total amounts) were multistationarity occurs. Moreover, in principle regions with more than 3 stoichiometrically compatible steady states can be detected. Our results have been implemented by Rick Rischter in a computer algebra system to find open regions in the space of all parameters where the 4-site and the 5-site distributive sequential phosphorylation cycles have at least 5 stoichiometrically compatible positive steady states[2].

Using Other Methods: As we remarked in the examples, the number of parameters is typically too high to try to naively use standard methods in computational algebraic geometry. We refer the reader to [**178**, **186**] for numerical methods and [**145**] for symbolic methods to detect sample points in each chamber of the

[2]Maple and SAGE files with these computations are available at: `http://mate.dm.uba.ar/~alidick/DGRPMFiles/`.

complement of the discriminant (and boundary resultant complexes) in the space of parameters. Also, it is shown in [**131**] that triangular forms can sometimes be used to detect multistationarity. Other authors explored the extrapolation of multistationarity from subnetworks in various ways (see for instance [**15**, **84**, **223**]) as the use of homotopy methods [**114**]. Researchers from other disciplines usually use numerical analysis and intensive parameter sampling.

Some Open Questions. We have already mentioned two concrete open questions on the maximum possible number of stoichiometrically compatible positive steady states in the distributive sequential n-site cycle and in the ERK pathway. In the first case, there is a conjecture for this maximum possible number $(2n-1)$, but it seems that all *degeneration* techniques to simplify the resulting system do not have more have $2\lfloor n/2 \rfloor + 1$ such positive steady states (as in [**150**, **168**]), which makes the problem complicated.

More generally, these are open questions in the study of sparse systems of polynomial equations with as many polynomials as variables: determine sharp upper and lower bounds for the number of positive solutions from the combinatorics of the supports, and from the interactions between the coefficients and the exponents of the system. This is only known when there is at most one root [**277**], an upper bound in the style of Descartes rule of signs for univariate polynomials is only known when the support of the polynomials consists of $m+2$ monomials in m variables [**29**], and in a few instances Khovanskii's bound has been tightly refined as in the case of two trinomials in two variables [**296**].

It is also important to be able to predict from the structure of the network the maximum possible number of stoichiometrically compatible positive steady states which are stable. Given a reaction network with known parameters and a known steady state, stability can be decided without computing the roots of the characteristic polynomial of the Jacobian by means of the classical Routh-Hurwitz criterion. But this is a hard task with steady states that can be only approximated (which is the standard situation) and moreover when dealing with parametric systems.

Another central question regarding the dynamics of biochemical reaction networks is to predict or preclude the ocurrence of oscillations from the structure. There are several recent articles with this approach [**14**, **82**, **85**, **200**, **284**], but this is mainly an open field. It is very interesting that even for the distributed sequential 2-site phosphorylation cycle it is not known whether it can have oscillatory solutions or not! Finally, once the capacity for oscillations has been determined, the next question would be to find oscillatory parameters and, better, oscillatory open regions in the space of parameters.

Final Comments. In this section we gave an overview of the state of the art and open questions in the algebro-geometric study of biochemical reaction networks modeling signaling pathways in systems biology. They give rise to quite diverse, but not general mass-action dynamical systems, which allows us to prove mathematical results and do computations based on their structure, even for systems with many variables and many parameters.

Illustration Credits

Page 6 Figure is from [**288**]. Created by Erwan Penchèvre. Used with permission of the creator.

Page 53 Top, middle and bottom figures are based on Figures 2 and 3 of [**343**]. Created by Simon Telen and Marc Van Barel. ©2018 Elsevier B.V. Reprinted with permission. All rights reserved.

Page 54 Top figure is based on Figure 3 of [**343**]. Created by Simon Telen and Marc Van Barel. ©2018 Elsevier B.V. Reprinted with permission. All rights reserved.

Page 54 Bottom figure is Figure 4 from [**13**]. Created by Henk Poulisse. Used with permission of the creator. Republished with permission of DeGruyter; permission conveyed through Copyright Clearance Center, Inc.

Page 63 Figure is Figure 2.1 from [**21**]. Created by Charles W. Wampler. ©2013 Society for Industrial and Applied Mathematics. Reprinted with permission. All rights reserved.

Page 94 Top figure is the left-most surface of Figure 12 from [**160**]. Created by Frank Sottile. Republished with permission of the Association for Computing Machinery; permission conveyed through Copyright Clearance Center, Inc.

Page 94 Bottom photo by Georges Jansoone. CC-BY-SA-3.0. `https://commons.wikimedia.org/wiki/File:Bilbao.Guggenheim08.jpg`

Page 95 Cover of SIAM Journal on Applied Algebra and Geometry ©2017 Society for Industrial and Applied Mathematics. Reprinted with permission. All rights reserved.

Page 98 Right figure in Example 3.4 is Figure 2 from [**327**]. Created by Frank Sottile. ©2003 American Mathematical Society. Reprinted with permission. All rights reserved.

Page 100 Upper-right figure of bottom display is the left-most surface of Figure 12 from [**160**]. Created by Frank Sottile. Republished with permission of the Association for Computing Machinery; permission conveyed through Copyright Clearance Center, Inc.

Page 101 Top figure is Figure 12 from [**160**]. Created by Frank Sottile. Republished with permission of the Association for Computing Machinery; permission conveyed through Copyright Clearance Center, Inc.

Page 137 Right figure in Example 4.2 created by Michael Ströck. CC-BY-SA-3.0 Mstroeck at the English Language Wikipedia and GNU Free Documentation license. `https://commons.wikimedia.org/wiki/File:Caffeine_Molecule.png`

Page 138 Figure in Example 4.4 is from [**35**]. Created by Oriol Bohigas, Montserrat Manubens and Lluís Ros. Republished with permission of American Society of Mechanical Engineers; Permission conveyed through Copyright Clearance Center, Inc.

Page 139 Figure in (4.1) is an image by Charles W. Wampler and Douglas N. Arnold. Used with permission of the creators.

Page 189 Figure in (5.20) created by Daniele Pugliesi. CC-BY-SA-3.0. `https://commons.wikimedia.org/wiki/File:Agitated_vessel.svg`

The images shown on the bottom of page 94, page 137 (right) and on page 189 are covered by the Creative Commons Attribution-Share Alike 3.0 license (CC-BY-SA-3.0):

`https://creativecommons.org/licenses/by-sa/3.0/deed.en`

The image shown on page 137 (right) is also licensed under the GNU Free Documentation license version 1.2 or later version.

Bibliography

[1] J. Abbott, A. Bigatti, *CoCoALib: a C++ library for doing Computations in Commutative Algebra*, `http://cocoa.dima.unige.it/cocoalib`.

[2] J. Abbott, A. M. Bigatti, and L. Robbiano, *Implicitization of hypersurfaces*, J. Symbolic Comput. **81** (2017), 20–40, DOI 10.1016/j.jsc.2016.11.002. MR3594322

[3] S. Abhyankar, *Polynomials and power series*, Math. Intelligencer **3**, Springer-Verlag, September 1972. Reprinted in *Algebra, Arithmetic and Geometry with Applications* (C. Christensen, G. Sundaram, A. Sathaye and C. Bajaj, eds.), Springer, New York, 2004, 783–784.

[4] S. S. Abhyankar, *Historical ramblings in algebraic geometry and related algebra*, Amer. Math. Monthly **83** (1976), no. 6, 409–448, DOI 10.2307/2318338. MR401754

[5] M. Adamer and M. Helmer, *Steady state invariants and multistationarity for families of toric reaction networks*, 2019, `arXiv:1906.03931[q-bio.MN]`.

[6] M. Aigner and G. M. Ziegler, *Proofs from The Book*, 5th ed., Springer-Verlag, Berlin, 2014. Including illustrations by Karl H. Hofmann. MR3288091

[7] H. Alt, *Über die Erzeugung gegebener ebener Kurven mit Hilfe des Gelenkvierseits*, Z. angew. Math. Mech. **3** (1923), 13–19.

[8] D. Arnold and C. Wampler, poster image created for the IMA thematic year on *IMA Applications of Algebraic Geometry*, `https://www.ima.umn.edu/2006-2007`.

[9] B. Assarf, E. Gawrilow, K. Herr, M. Joswig, B. Lorenz, A. Paffenholz, and T. Rehn, *Computing convex hulls and counting integer points with* `polymake`, Math. Program. Comput. **9** (2017), no. 1, 1–38, DOI 10.1007/s12532-016-0104-z. MR3613012

[10] L. Asimow and B. Roth, *The rigidity of graphs*, Trans. Amer. Math. Soc. **245** (1978), 279–289, DOI 10.2307/1998867. MR511410

[11] L. Asimow and B. Roth, *The rigidity of graphs. II*, J. Math. Anal. Appl. **68** (1979), no. 1, 171–190, DOI 10.1016/0022-247X(79)90108-2. MR531431

[12] W. Auzinger and H. J. Stetter, *An elimination algorithm for the computation of all zeros of a system of multivariate polynomial equations*, Numerical mathematics, Singapore 1988, Internat. Schriftenreihe Numer. Math., vol. 86, Birkhäuser, Basel, 1988, pp. 11–30. MR1022943

[13] C. Baciu and M. Kreuzer, *Algebraisches Öl*, Mitteilungen der DMV **19** (2011), 142–147.

[14] M. Banaji, *Inheritance of oscillation in chemical reaction networks*, Appl. Math. Comput. **325** (2018), 191–209, DOI 10.1016/j.amc.2017.12.012. MR3759136

[15] M. Banaji and C. Pantea, *The inheritance of nondegenerate multistationarity in chemical reaction networks*, SIAM J. Appl. Math. **78** (2018), no. 2, 1105–1130, DOI 10.1137/16M1103506. MR3784124

[16] E. Bartzos, I. Emiris, J. Legerský and E. Tsigaridas, *On the maximal number of real embeddings of minimally rigid graphs in* $\mathbb{R}^2$, $\mathbb{R}^3$ *and* S^2, J. Symbolic Comput. (2019), to appear, DOI 10.1016/j.jsc.2019.10.015.

[17] I. G. Bashmakova and G. S. Smirnova, *The beginnings and evolution of algebra*, The Dolciani Mathematical Expositions, vol. 23, Mathematical Association of America, Washington, DC, 2000. Translated from the Russian by Abe Shenitzer with the editorial assistance of David A. Cox. MR1735499

[18] A. Baskar and S. Bandyopadhyay, *An algorithm to compute the finite roots of large systems of polynomial equations arising in kinematic synthesis*, Mech. Mach. Theory **133** (2019), 493–513.

[19] D. J. Bates, J. D. Hauenstein, A. J. Sommese, and C. W. Wampler II, *Adaptive multiprecision path tracking*, SIAM J. Numer. Anal. **46** (2008), no. 2, 722–746, DOI 10.1137/060658862. MR2383209

[20] D. Bates, J. Hauenstein, A. Sommese and C. Wampler, *Bertini: Software for Numerical Algebraic Geometry*, `https://bertini.nd.edu`.

[21] D. J. Bates, J. D. Hauenstein, A. J. Sommese, and C. W. Wampler, *Numerically solving polynomial systems with Bertini*, Software, Environments, and Tools, vol. 25, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2013. MR3155500

[22] D. J. Bates, J. D. Hauenstein, T. M. McCoy, C. Peterson, and A. J. Sommese, *Recovering exact results from inexact numerical data in algebraic geometry*, Exp. Math. **22** (2013), no. 1, 38–50, DOI 10.1080/10586458.2013.737640. MR3038781

[23] C. Beltrán and L. M. Pardo, *On Smale's 17th problem: a probabilistic positive solution*, Found. Comput. Math. **8** (2008), no. 1, 1–43, DOI 10.1007/s10208-005-0211-0. MR2403529

[24] C. Beltrán and L. M. Pardo, *Smale's 17th problem: average polynomial time to compute affine and projective solutions*, J. Amer. Math. Soc. **22** (2009), no. 2, 363–385, DOI 10.1090/S0894-0347-08-00630-9. MR2476778

[25] D. N. Bernstein, *The number of roots of a system of equations* (Russian), Funkcional. Anal. i Priložen. **9** (1975), no. 3, 1–4. English translation in Functional Anal. Appl. **9** (1975), 183–185. MR0435072

[26] S. Bernstein, *Démonstration du théorème de Weierstrass fondée sur le calcul des probabilités*, Comm. Kharkov Math. Soc. **13** (1912/13), 1–2.

[27] E. Bézout, *Sur le degré des équations résultantes de l'évanouissement des inconnues*, Histoire de l'Académie Royale des Sciences (1764), 288–338.

[28] E. Bézout, *General theory of algebraic equations*, Princeton University Press, Princeton, NJ, 2006. Translated from the 1779 French original by Eric Feron. MR2215193

[29] F. Bihan and A. Dickenstein, *Descartes' rule of signs for polynomial systems supported on circuits*, Int. Math. Res. Not. IMRN **22** (2017), 6867–6893, DOI 10.1093/imrn/rnw199. MR3737324

[30] F. Bihan, A. Dickenstein and M. Giaroli, *Lower bounds for positive roots and regions of multistationarity in chemical reaction networks*, 2018, `arXiv:1807.05157[math.DS]`.

[31] F. Bihan, J. M. Rojas, and F. Sottile, *On the sharpness of fewnomial bounds and the number of components of fewnomial hypersurfaces*, Algorithms in algebraic geometry, IMA Vol. Math. Appl., vol. 146, Springer, New York, 2008, pp. 15–20, DOI 10.1007/978-0-387-75155-9_2. MR2397935

[32] F. Bihan, F. Santos, and P.-J. Spaenlehauer, *A polyhedral method for sparse systems with many positive solutions*, SIAM J. Appl. Algebra Geom. **2** (2018), no. 4, 620–645, DOI 10.1137/18M1181912. MR3892410

[33] L. Blum, F. Cucker, M. Shub, and S. Smale, *Complexity and real computation*, Springer-Verlag, New York, 1998. With a foreword by Richard M. Karp. MR1479636

[34] L. M. Blumenthal, *Theory and applications of distance geometry*, Oxford, at the Clarendon Press, 1953. MR0054981

[35] O. Bohigas, M. Manubens, and L. Ros, *A linear relaxation method for computing workspace slices of the Stewart platform*, J. Mechanisms Robotics **5** (2013), issue 1.

[36] C. Borcea and I. Streinu, *The number of embeddings of minimally rigid graphs*, Discrete Comput. Geom. **31** (2004), no. 2, 287–303, DOI 10.1007/s00454-003-2902-0. MR2060642

[37] B. Boros, *Existence of positive steady states for weakly reversible mass-action systems*, SIAM J. Math. Anal. **51** (2019), no. 1, 435–449, DOI 10.1137/17M115534X. MR3911717

[38] J. A. Boswell and V. Mukundan, *Rees algebras and almost linearly presented ideals*, J. Algebra **460** (2016), 102–127, DOI 10.1016/j.jalgebra.2016.03.035. MR3510395

[39] N. Botbol, *The implicit equation of a multigraded hypersurface*, J. Algebra **348** (2011), 381–401, DOI 10.1016/j.jalgebra.2011.09.019. MR2852248

[40] N. Botbol, L. Busé, and M. Chardin, *Fitting ideals and multiple points of surface parameterizations*, J. Algebra **420** (2014), 486–508, DOI 10.1016/j.jalgebra.2014.07.028. MR3261469

[41] N. Botbol and A. Dickenstein, *Implicitization of rational hypersurfaces via linear syzygies: a practical overview*, J. Symbolic Comput. **74** (2016), 493–512, DOI 10.1016/j.jsc.2015.09.001. MR3424053

[42] N. Botbol, A. Dickenstein, and M. Dohm, *Matrix representations for toric parametrizations*, Comput. Aided Geom. Design **26** (2009), no. 7, 757–771, DOI 10.1016/j.cagd.2009.03.005. MR2569833

[43] A. Brill, *Ueber eine Eigenschaft der Resultante* (German), Math. Ann. **16** (1880), no. 3, 345–347, DOI 10.1007/BF01447121. MR1510029

[44] A. Brill, *Elimination und Geometrie in den letzten Jahrzehnten*, in *Verhandlungen des dritten internationalen Mathematiker-Kongresses in Heidelberg 1904*, Leipzig 1905, 275–283.

[45] A. Brill and M. Noether, *Die Entwicklung der Theorie der algebraischen Functionen in älterer und neuerer Zeit*, Jahresbericht der Deutschen Mathematiker Vereinigung **3** (1892–1893), 111–566,

[46] E. Brugallé, I. Itenberg, G. Mikhalkin and K. Shaw, *Brief introduction to tropical geometry*, in *Proceedings of the Gökova* (S. Akbulut, D. Auroux and T. Önder, eds.), International Press, Boston, 2016, 1–75.

[47] B. Buchberger, *Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem nulldimensionalen Polynomideal*, Ph.D. Thesis, University of Innsbruck, 1965.

[48] B. Buchberger, *Ein algorithmisches Kriterium für die Lösbarkeit eines algebraischen Gleichungssystems* (German), Aequationes Math. **4** (1970), 374–383, DOI 10.1007/BF01844169. English translation *An algorithmic criterion for the solvability of a system of algebraic equations* by M. Abramson and R. Lumbert in *Gröbner Bases and Applications* (B. Buchberger and F. Winkler, eds.) MR0268178

[49] B. Buchberger and H. M. Möller, *The construction of multivariate polynomials with preassigned zeros*, Computer algebra (Marseille, 1982), Lecture Notes in Comput. Sci., vol. 144, Springer, Berlin-New York, 1982, pp. 24–31. MR680050

[50] P. Bürgisser and F. Cucker, *Condition: The geometry of numerical algorithms*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 349, Springer, Heidelberg, 2013. MR3098452

[51] P. Bürgisser and F. Cucker, *On a problem posed by Steve Smale*, Ann. of Math. (2) **174** (2011), no. 3, 1785–1836, DOI 10.4007/annals.2011.174.3.8. MR2846491

[52] L. Busé, *Residual resultant over the projective plane and the implicitization problem*, Proceedings of the 2001 International Symposium on Symbolic and Algebraic Computation, ACM, New York, 2001, pp. 48–55, DOI 10.1145/384101.384109. MR2049730

[53] L. Busé, *Resultants of determinantal varieties*, J. Pure Appl. Algebra **193** (2004), no. 1-3, 71–97, DOI 10.1016/j.jpaa.2004.02.010. MR2076379

[54] L. Busé, *On the equations of the moving curve ideal of a rational algebraic plane curve*, J. Algebra **321** (2009), no. 8, 2317–2344, DOI 10.1016/j.jalgebra.2009.01.030. MR2501523

[55] B. Laurent, *Implicit matrix representations of rational Bézier curves and surfaces*, Comput.-Aided Des. **46** (2014), 14–24, DOI 10.1016/j.cad.2013.08.014. MR3124172

[56] L. Busé and T. Luu Ba, *The surface/surface intersection problem by means of matrix based representations*, Comput. Aided Geom. Design **29** (2012), no. 8, 579–598, DOI 10.1016/j.cagd.2012.04.002. MR2948569

[57] L. Busé and M. Chardin, *Implicitizing rational hypersurfaces using approximation complexes*, J. Symbolic Comput. **40** (2005), no. 4-5, 1150–1168, DOI 10.1016/j.jsc.2004.04.005. MR2172855

[58] L. Busé, M. Chardin, and J.-P. Jouanolou, *Torsion of the symmetric algebra and implicitization*, Proc. Amer. Math. Soc. **137** (2009), no. 6, 1855–1865, DOI 10.1090/S0002-9939-09-09550-1. MR2480264

[59] L. Busé, M. Chardin, and A. Simis, *Elimination and nonlinear equations of Rees algebras*, J. Algebra **324** (2010), no. 6, 1314–1333, DOI 10.1016/j.jalgebra.2010.07.006. With an appendix in French by Joseph Oesterlé. MR2671807

[60] L. Busé, D. Cox, and C. D'Andrea, *Implicitization of surfaces in $\mathbb{P}^3$ in the presence of base points*, J. Algebra Appl. **2** (2003), no. 2, 189–214, DOI 10.1142/S0219498803000489. MR1980408

[61] L. Busé and C. D'Andrea, *Singular factors of rational plane curves*, J. Algebra **357** (2012), 322–346, DOI 10.1016/j.jalgebra.2012.01.030. MR2905259

[62] L. Busé, M. Elkadi, and B. Mourrain, *Generalized resultants over unirational algebraic varieties*, J. Symbolic Comput. **29** (2000), no. 4-5, 515–526, DOI 10.1006/jsco.1999.0304. Symbolic computation in algebra, analysis, and geometry (Berkeley, CA, 1998). MR1769653

[63] L. Busé, M. Elkadi, and B. Mourrain, *Resultant over the residual of a complete intersection*, J. Pure Appl. Algebra **164** (2001), no. 1-2, 35–57, DOI 10.1016/S0022-4049(00)00144-4. Effective methods in algebraic geometry (Bath, 2000). MR1854329

[64] L. Busé and J.-P. Jouanolou, *On the closed image of a rational map and the implicitization problem*, J. Algebra **265** (2003), no. 1, 312–357, DOI 10.1016/S0021-8693(03)00181-9. MR1984914

[65] L. Busé and A. Karasoulou, *Resultant of an equivariant polynomial system with respect to the symmetric group*, J. Symbolic Comput. **76** (2016), 142–157, DOI 10.1016/j.jsc.2015.12.004. MR3461263

[66] J. Capco, M. Gallet, G. Grasegger, C. Koutschan, N. Lubbes, and J. Schicho, *The number of realizations of a Laman graph*, SIAM J. Appl. Algebra Geom. **2** (2018), no. 1, 94–125, DOI 10.1137/17M1118312. A brief version was published as [**67**]. MR3771397

[67] J. Capco, M. Gallet, G. Grasegger, C. Koutschan, N. Lubbes and J. Schicho, *Computing the number of realizations of a Laman graph*, Electronic Notes in Disc. Math. **61** (2017), 207–213. See [**66**] for more details.

[68] A. Cauchy, *Recherches sur les polyèdres, Premier Mémoire*, Journal de l'École Polytechnique **9** (1813), 68–86.

[69] A. Cayley, *A theorem in the geometry of position*, Cambridge Math. Journal **2** (1841), 267–271.

[70] A. Cayley, *On linear transformations*, Cambridge and Dublin Math. Journal **1** (1846), 104–122.

[71] A. Cayley, *Recherches sur l'élimination, et sur la théorie des courbes* (French), J. Reine Angew. Math. **34** (1847), 30–45, DOI 10.1515/crll.1847.34.30. MR1578563

[72] A. Cayley, *Note sur la méthode d'élimination de Bezout* (French), J. Reine Angew. Math. **53** (1857), 366–367, DOI 10.1515/crll.1857.53.366. MR1579014

[73] A. Cayley, *A fourth memoir upon quantics*, Phil. Trans. **148** (1858), 415–427.

[74] A. Cayley, *Nouvelles recherches sur l'élimination et la théorie des courbes* (French), J. Reine Angew. Math. **63** (1864), 34–39, DOI 10.1515/crll.1864.63.34. MR1579252

[75] M. Chardin, *Implicitization using approximation complexes*, Algebraic geometry and geometric modeling, Math. Vis., Springer, Berlin, 2006, pp. 23–35, DOI 10.1007/978-3-540-33275-6_2. MR2279841

[76] F. Chen, D. Cox, and Y. Liu, *The μ-basis and implicitization of a rational parametric surface*, J. Symbolic Comput. **39** (2005), no. 6, 689–706, DOI 10.1016/j.jsc.2005.01.003. MR2168614

[77] F. Chen, L. Shen, and J. Deng, *Implicitization and parametrization of quadratic and cubic surfaces by μ-bases*, Computing **79** (2007), no. 2-4, 131–142, DOI 10.1007/s00607-006-0192-0. MR2295516

[78] T. Chen, T.-L. Lee, and T.-Y. Li, *Hom4PS-3: a parallel numerical solver for systems of polynomial equations based on polyhedral homotopy continuation methods*, Mathematical software—ICMS 2014, Lecture Notes in Comput. Sci., vol. 8592, Springer, Heidelberg, 2014, pp. 183–190, DOI 10.1007/978-3-662-44199-2_30. MR3334764

[79] B. Clarke, *Stability of complex reaction networks*, Adv. Chem. Phys. **42** (1980), 1–213.

[80] C. Clement, A. Lee-St. John, and J. Sidman, *Hyperbanana graphs*, Proceedings of 25th Canadian Conference on Computational Geometry, (2013), paper 45, `http://cccg.ca/proceedings/2013/`.

[81] D. Cifuentes and P. A. Parrilo, *Sampling algebraic varieties for sum of squares programs*, SIAM J. Optim. **27** (2017), no. 4, 2381–2404, DOI 10.1137/15M1052548. MR3724261

[82] C. Conradi, E. Feliu and M. Mincheva, *On the existence of Hopf bifurcations in the sequential and distributive double phosphorylation cycle*, 2019, `arXiv:1905.08129[q-bio.MN]`.

[83] C. Conradi, E. Feliu, M. Mincheva and C. Wiuf, *Identifying parameter regions for multistationarity*, PLoS Comput. Biol. **13** (2017), e1005751.

[84] C. Conradi, D. Flockerzi, J. Raisch and J. Stelling, *Subnetwork analysis reveals dynamic features of complex (bio)chemical networks*, PNAS **104** (2007), 19175–19180.

[85] C. Conradi, M. Mincheva, and A. Shiu, *Emergence of oscillations in a mixed-mechanism phosphorylation system*, Bull. Math. Biol. **81** (2019), no. 6, 1829–1852, DOI 10.1007/s11538-019-00580-6. MR3945912

[86] C. Conradi and C. Pantea, *Multistationarity in biochemical networks: results, analysis, and examples*, Algebraic and combinatorial computational biology, Math. Sci. Eng., Academic Press, London, 2019, pp. 279–317. MR3839594

[87] C. Conradi and A. Shiu, *A global convergence result for processive multisite phosphorylation systems*, Bull. Math. Biol. **77** (2015), no. 1, 126–155, DOI 10.1007/s11538-014-0054-4. MR3303108

[88] C. Conradi and A. Shiu, *Dynamics of post-translational modification systems: recent progress and future directions*, Biophys. J. **114** (2018), 507–515.

[89] T. Cortadellas Benítez and C. D'Andrea, *Minimal generators of the defining ideal of the Rees algebra associated to monoid parameterizations*, Comput. Aided Geom. Design **27** (2010), no. 6, 461–473, DOI 10.1016/j.cagd.2010.04.003. MR2657547

[90] T. Cortadellas Benítez and C. D'Andrea, *Minimal generators of the defining ideal of the Rees algebra associated with a rational plane parametrization with* $\mu = 2$, Canad. J. Math. **66** (2014), no. 6, 1225–1249, DOI 10.4153/CJM-2013-035-1. MR3270782

[91] T. Cortadellas Benítez and C. D'Andrea, *Rational plane curves parameterizable by conics*, J. Algebra **373** (2013), 453–480, DOI 10.1016/j.jalgebra.2012.09.034. MR2995038

[92] T. Cortadellas Benítez, D. A. Cox, and C. D'Andrea, *The Rees algebra of parametric curves via liftings*, J. Pure Appl. Algebra **224** (2020), no. 2, 869–893, DOI 10.1016/j.jpaa.2019.06.015. MR3987981

[93] M. Coste, *An Introduction to Semialgebraic Geometry*, Univ. of Rennes, 2002, `http://gcomte.perso.math.cnrs.fr/M2/CosteIntroToSemialGeo.pdf`.

[94] D. A. Cox, *Equations of parametric curves and surfaces via syzygies*, Symbolic computation: solving equations in algebra, geometry, and engineering (South Hadley, MA, 2000), Contemp. Math., vol. 286, Amer. Math. Soc., Providence, RI, 2001, pp. 1–20, DOI 10.1090/conm/286/04751. MR1874268

[95] D. Cox, *Curves, surfaces, and syzygies*, Topics in algebraic geometry and geometric modeling, Contemp. Math., vol. 334, Amer. Math. Soc., Providence, RI, 2003, pp. 131–150, DOI 10.1090/conm/334/05979. MR2039970

[96] D. A. Cox, *The moving curve ideal and the Rees algebra*, Theoret. Comput. Sci. **392** (2008), no. 1-3, 23–36, DOI 10.1016/j.tcs.2007.10.012. MR2394983

[97] D. A. Cox, *Solving equations via algebras*, Solving polynomial equations, Algorithms Comput. Math., vol. 14, Springer, Berlin, 2005, pp. 63–123, DOI 10.1007/3-540-27357-3_2. MR2161986

[98] D. Cox, *What is a toric variety?*, Topics in algebraic geometry and geometric modeling, Contemp. Math., vol. 334, Amer. Math. Soc., Providence, RI, 2003, pp. 203–223, DOI 10.1090/conm/334/05983. MR2039974

[99] D. Cox and P. Clarke, *Moment maps, strict linear precision, and maximum likelihood degree one*, 2018, `arXiv:1810.03672[math.AG]`.

[100] D. Cox, R. Goldman, and M. Zhang, *On the validity of implicitization by moving quadrics of rational surfaces with no base points*, J. Symbolic Comput. **29** (2000), no. 3, 419–440, DOI 10.1006/jsco.1999.0325. MR1751389

[101] D. Cox, J. W. Hoffman, and H. Wang, *Syzygies and the Rees algebra*, J. Pure Appl. Algebra **212** (2008), no. 7, 1787–1796, DOI 10.1016/j.jpaa.2007.11.006. MR2400743

[102] D. A. Cox and A. A. Iarrobino, *Strata of rational space curves*, Comput. Aided Geom. Design **32** (2015), 50–68, DOI 10.1016/j.cagd.2014.11.004. MR3301274

[103] D. Cox, A. R. Kustin, C. Polini, and B. Ulrich, *A study of singularities on rational curves via syzygies*, Mem. Amer. Math. Soc. **222** (2013), no. 1045, x+116, DOI 10.1090/S0065-9266-2012-00674-5. MR3059370

[104] D. A. Cox, J. Little, and D. O'Shea, *Ideals, varieties, and algorithms*: *An introduction to computational algebraic geometry and commutative algebra*, 4th ed., Undergraduate Texts in Mathematics, Springer, Cham, 2015. MR3330490

[105] D. A. Cox, J. Little, and D. O'Shea, *Using algebraic geometry*, 2nd ed., Graduate Texts in Mathematics, vol. 185, Springer, New York, 2005. MR2122859

[106] D. A. Cox, J. B. Little, and H. K. Schenck, *Toric varieties*, Graduate Studies in Mathematics, vol. 124, American Mathematical Society, Providence, RI, 2011. MR2810322

[107] D. Cox, K.-N. Lin and G. Sosa, *Multi-Rees algebras and toric dynamical systems*, Proc. Amer. Math. Soc., to appear.

[108] D. A. Cox, T. W. Sederberg, and F. Chen, *The moving line ideal basis of planar rational curves*, Comput. Aided Geom. Design **15** (1998), no. 8, 803–827, DOI 10.1016/S0167-8396(98)00014-4. MR1638732

[109] G. Craciun, *Toric differential inclusions and a proof of the global attractor conjecture*, 2015, arXiv:1501.02860[math.DS].

[110] G. Craciun, *Polynomial dynamical systems, reaction networks, and toric differential inclusions*, SIAM J. Appl. Algebra Geom. **3** (2019), no. 1, 87–106, DOI 10.1137/17M1129076. MR3920470

[111] G. Craciun and M. Feinberg, *Multiple equilibria in complex chemical reaction networks. I. The injectivity property*, SIAM J. Appl. Math. **65** (2005), no. 5, 1526–1546, DOI 10.1137/S0036139904440278. MR2177713

[112] G. Craciun and M. Feinberg, *Multiple equilibria in complex chemical reaction networks: semiopen mass action systems*, SIAM J. Appl. Math. **70** (2010), no. 6, 1859–1877, DOI 10.1137/090756387. MR2596505

[113] G. Craciun, A. Dickenstein, A. Shiu, and B. Sturmfels, *Toric dynamical systems*, J. Symbolic Comput. **44** (2009), no. 11, 1551–1565, DOI 10.1016/j.jsc.2008.08.006. MR2561288

[114] G. Craciun, J. W. Helton, and R. J. Williams, *Homotopy methods for counting reaction network equilibria*, Math. Biosci. **216** (2008), no. 2, 140–149, DOI 10.1016/j.mbs.2008.09.001. MR2477000

[115] G. Craciun, L. D. García-Puente, and F. Sottile, *Some geometrical aspects of control points for toric patches*, Mathematical methods for curves and surfaces, Lecture Notes in Comput. Sci., vol. 5862, Springer, Berlin, 2010, pp. 111–135, DOI 10.1007/978-3-642-11620-9_9. MR3193313

[116] G. Cramer, *Introduction à l'analyse des lignes courbes algébriques*, Frères Cramer et Cl. Philibert, Genêve, 1750.

[117] N. S. Daleo and J. D. Hauenstein, *Numerically deciding the arithmetically Cohen-Macaulayness of a projective scheme*, J. Symbolic Comput. **72** (2016), 128–146, DOI 10.1016/j.jsc.2015.01.001. MR3369053

[118] C. D'Andrea, H. Hong, T. Krick, and A. Szanto, *An elementary proof of Sylvester's double sums for subresultants*, J. Symbolic Comput. **42** (2007), no. 3, 290–297, DOI 10.1016/j.jsc.2006.09.003. MR2292636

[119] C. D'Andrea and A. Khetan, *Implicitization of rational surfaces using toric varieties*, J. Algebra **303** (2006), no. 2, 543–565, DOI 10.1016/j.jalgebra.2005.05.028. MR2255122

[120] C. D'Andrea and M. Sombra, *A Poisson formula for the sparse resultant*, Proc. Lond. Math. Soc. (3) **110** (2015), no. 4, 932–964, DOI 10.1112/plms/pdu069. MR3335291

[121] W. Decker, G-M. Greuel, G. Pfister and H. Schönemann, SINGULAR *(4-1-1) — A computer algebra system for polynomial computations*, 2018, http://www.singular.uni-kl.de.

[122] W. L. F. Degen, *The types of rational* $(2,1)$*-Bézier surfaces*, Comput. Aided Geom. Design **16** (1999), no. 7, 639–648, DOI 10.1016/S0167-8396(99)00028-X. Dedicated to Paul de Faget de Casteljau. MR1718055

[123] P. De Leenheer, *An elementary proof of a matrix tree theorem for directed graphs*, 2019, arXiv:1904.12221[math.CO].

[124] P. Deuflhard and A. Hohmann, *Numerical analysis in modern scientific computing*: *An introduction*, 2nd ed., Texts in Applied Mathematics, vol. 43, Springer-Verlag, New York, 2003. MR1949263

[125] S. Di Rocco, D. Eklund, C. Peterson, and A. J. Sommese, *Chern numbers of smooth varieties via homotopy continuation and intersection theory*, J. Symbolic Comput. **46** (2011), no. 1, 23–33, DOI 10.1016/j.jsc.2010.06.026. MR2736356

[126] A. Dickenstein, *Algebra and geometry in the study of enzymatic cascades*, in *World Women in Mathematics 2018 – Proceedings of the First World Meeting for Women in Mathematics* $(WM)^2$ (C. Araujo, G. Benkart, C. Praeger and B. Tanbay, eds.), Association for Women in Mathematics Series, Springer, 2019.

[127] A. Dickenstein, *Biochemical reaction networks: an invitation for algebraic geometers*, Mathematical Congress of the Americas, Contemp. Math., vol. 656, Amer. Math. Soc., Providence, RI, 2016, pp. 65–83, DOI 10.1090/conm/656/13076. MR3457596

[128] A. Dickenstein and I. Z. Emiris, *Multihomogeneous resultant formulae by means of complexes*, J. Symbolic Comput. **36** (2003), no. 3-4, 317–342, DOI 10.1016/S0747-7171(03)00086-5. International Symposium on Symbolic and Algebraic Computation (ISSAC'2002) (Lille). MR2004032

[129] A. Dickenstein and E. Feliu, *Algebraic Methods for Biochemical Reaction Networks*, in preparation.

[130] A. Dickenstein, E. M. Feichtner, and B. Sturmfels, *Tropical discriminants*, J. Amer. Math. Soc. **20** (2007), no. 4, 1111–1133, DOI 10.1090/S0894-0347-07-00562-0. MR2328718

[131] A. Dickenstein, M. Pérez Millán, A. Shiu, and X. Tang, *Multistationarity in structured reaction networks*, Bull. Math. Biol. **81** (2019), no. 5, 1527–1581, DOI 10.1007/s11538-019-00572-6. MR3935185

[132] R. Diestel, *Graph theory*, 5th ed., Graduate Texts in Mathematics, vol. 173, Springer, Berlin, 2017. MR3644391

[133] A. L. Dixon, *The eliminant of three quantics in two independent variables*, Proc. London Math. Soc. (2) **7** (1909), 49–69, DOI 10.1112/plms/s2-7.1.49. MR1575687

[134] E. Duarte and H. Schenck, *Tensor product surfaces and linear syzygies*, Proc. Amer. Math. Soc. **144** (2016), no. 1, 65–72, DOI 10.1090/proc/12703. MR3415577

[135] E. Dufresne, P. Edwards, H. Harrington and J. Hauenstein, *Sampling real algebraic varieties for topological data analysis*, `arXiv:1802.07716[math.AT]`.

[136] J. Draisma, E. Horobeţ, G. Ottaviani, B. Sturmfels, and R. R. Thomas, *The Euclidean distance degree of an algebraic variety*, Found. Comput. Math. **16** (2016), no. 1, 99–149, DOI 10.1007/s10208-014-9240-x. MR3451425

[137] B. Edelstein, *Biochemical models with multiple steady states and hysteresis*, J. Theor. Biol. **29** (1970), 57–62.

[138] W. Edge, *The theory of ruled surfaces*, Cambridge Univ, Press, 1931.

[139] D. Eisenbud, *Commutative algebra*: *With a view toward algebraic geometry*, Graduate Texts in Mathematics, vol. 150, Springer-Verlag, New York, 1995. MR1322960

[140] D. Eisenbud, *The ReesAlgebra package in Macaulay2*, J. Softw. Algebra Geom. **8** (2018), 49–60, DOI 10.2140/jsag.2018.8.49. MR3857649

[141] D. Eisenbud and B. Sturmfels, *Binomial ideals*, Duke Math. J. **84** (1996), no. 1, 1–45, DOI 10.1215/S0012-7094-96-08401-X. MR1394747

[142] M. Elkadi, A. Galligo, and T. H. Lê, *Parametrized surfaces in* $\mathbb{P}^3$ *of bidegree* $(1, 2)$, ISSAC 2004, ACM, New York, 2004, pp. 141–148, DOI 10.1145/1005285.1005307. MR2126936

[143] F. Enriques and O. Chisini, *Lezioni sulla teoria geometrica delle equazioni e delle funzioni algebriche. Vol. I* (Italian), Collana di Matematica [Mathematics Collection], vol. 5, Zanichelli Editore S.p.A., Bologna, 1985. Reprint of the 1915 edition. MR966664

[144] J. Farre, H. Kleinschmidt, J. Sidman, A. St. John, S. Stark, L. Theran, and X. Yu, *Algorithms for detecting dependencies and rigid subsystems for CAD*, Comput. Aided Geom. Design **47** (2016), 130–149, DOI 10.1016/j.cagd.2016.06.001. MR3545987

[145] J.-C. Faugère, G. Moroz, F. Rouillier, and M. Safey El Din, *Classification of the perspective-three-point problem, discriminant variety and real solving polynomial systems of inequalities*, ISSAC 2008, ACM, New York, 2008, pp. 79–86, DOI 10.1145/1390768.1390782. MR2500376

[146] I. D. Faux and M. J. Pratt, *Computational geometry for design and manufacture*, Ellis Horwood Ltd., Chichester; distributed by Halsted Press [John Wiley & Sons], New York-Chichester-Brisbane, 1979. Mathematics and its Applications. MR524724

[147] M. Feinberg, *Complex balancing in general kinetic systems*, Arch. Rational Mech. Anal. **49** (1972/73), 187–194, DOI 10.1007/BF00255665. MR413930

[148] M. Feinberg, *Lectures On Chemical Reaction Networks*, 1979, `https://crnt.osu.edu/LecturesOnReactionNetworks`.

[149] M. Feinberg and F. J. M. Horn, *Chemical mechanism structure and the coincidence of the stoichiometric and kinetic subspaces*, Arch. Rational. Mech. Anal. **66** (1977), no. 1, 83–97, DOI 10.1007/BF00250853. MR0446109

[150] E. Feliu, A. Rendall and C. Wiuf, *A proof of unlimited multistability for phosphorylation cycles*, 2019, `arXiv:1904.02983[q-bio.MN]`.

[151] E. Feliu and C. Wiuf, *Variable elimination in chemical reaction networks with mass-action kinetics*, SIAM J. Appl. Math. **72** (2012), no. 4, 959–981, DOI 10.1137/110847305. MR2968758

[152] E. Feliu and C. Wiuf, *Enzyme-sharing as a cause of multi-stationarity in signalling systems*, J. R. Soc. Interface **9** (2011), 1224–1232.

[153] E. Feliu and C. Wiuf, *Simplifying biochemical models with intermediate species*, J. R. Soc. Interface **10** (2013), 20130484.

[154] H. Ferguson and D. Bailey, *A polynomial time, numerically stable integer relation algorithm*, RNR Techn. Rept. (1992), RNR-91-032.

[155] D. Flockerzi, K. Holstein, and C. Conradi, *n-site phosphorylation systems with* $2n-1$ *steady states*, Bull. Math. Biol. **76** (2014), no. 8, 1892–1916, DOI 10.1007/s11538-014-9984-0. MR3255149

[156] L. Fouli and K.-N. Lin, *Rees algebras of square-free monomial ideals*, J. Commut. Algebra **7** (2015), no. 1, 25–54, DOI 10.1216/JCA-2015-7-1-25. MR3316984

[157] W. Fulton, *Introduction to toric varieties*, Annals of Mathematics Studies, vol. 131, Princeton University Press, Princeton, NJ, 1993. The William H. Roever Lectures in Geometry. MR1234037

[158] A. Galligo and T.-H. Lê, *General classification of* $(1,2)$ *parametric surfaces in* $\mathbb{P}^3$, Geometric modeling and algebraic geometry, Springer, Berlin, 2008, pp. 93–113, DOI 10.1007/978-3-540-72185-7_6. MR2381606

[159] L. D. Garcia-Puente and F. Sottile, *Linear precision for parametric patches*, Adv. Comput. Math. **33** (2010), no. 2, 191–214, DOI 10.1007/s10444-009-9126-7. MR2659586

[160] L. García-Puente, F. Sottile and C. Zhu, *Toric degenerations of Bézier patches*, ACM Transactions on Graphics **30** (2011), 110:1–110:10.

[161] K. Gatermann, *Counting stable solutions of sparse polynomial systems in chemistry*, Symbolic computation: solving equations in algebra, geometry, and engineering (South Hadley, MA, 2000), Contemp. Math., vol. 286, Amer. Math. Soc., Providence, RI, 2001, pp. 53–69, DOI 10.1090/conm/286/04754. MR1874271

[162] K. Gatermann and B. Huber, *A family of sparse polynomial systems arising in chemical reaction systems*, J. Symbolic Comput. **33** (2002), no. 3, 275–305, DOI 10.1006/jsco.2001.0512. MR1882230

[163] K. Gatermann and M. Wolfrum, *Bernstein's second theorem and Viro's method for sparse polynomial systems in chemistry*, Adv. in Appl. Math. **34** (2005), no. 2, 252–294, DOI 10.1016/j.aam.2004.04.003. MR2110552

[164] K. Gatermann, M. Eiswirth, and A. Sensse, *Toric ideals and graph theory to analyze Hopf bifurcations in mass action systems*, J. Symbolic Comput. **40** (2005), no. 6, 1361–1382, DOI 10.1016/j.jsc.2005.07.002. MR2178092

[165] I. M. Gel′fand, M. M. Kapranov, and A. V. Zelevinsky, *Discriminants, resultants, and multidimensional determinants*, Mathematics: Theory & Applications, Birkhäuser Boston, Inc., Boston, MA, 1994. MR1264417

[166] G. Giambelli, *Sulle varietà rappresentate coll'annullare determinanti minori contenuti in un determinante simmetrico od emisimmetrico generico di forme*, Atti R. Accad. Sci. Torino **44** (1905/1906), 102–125.

[167] M. Giaroli, F. Bihan, and A. Dickenstein, *Regions of multistationarity in cascades of Goldbeter-Koshland loops*, J. Math. Biol. **78** (2019), no. 4, 1115–1145, DOI 10.1007/s00285-018-1304-0. MR3922967

[168] M. Giaroli, R. Rischter, M. Pérez Millán and A. Dickenstein, *Parameter regions that give rise to* $2\lfloor n/2\rfloor+1$ *positive steady states in the n-site phosphorylation system*, 2019, `arXiv:1904.11633[q-bio.MN]`.

[169] J. Giraud, *Géométrie algébrique élémentaire* (French), Département de Mathématique, Université de Paris-Sud, Orsay, 1977. Cours de troisième cycle (1975–76); Publications Mathématiques d'Orsay, No. 77-75. MR0491670

[170] M. Giusti, J. Heintz, and J. Sabia, *On the efficiency of effective Nullstellensätze*, Comput. Complexity **3** (1993), no. 1, 56–95, DOI 10.1007/BF01200407. MR1220078

[171] M. Giusti, J. Heintz, J. E. Morais, J. Morgenstern, and L. M. Pardo, *Straight-line programs in geometric elimination theory*, J. Pure Appl. Algebra **124** (1998), no. 1-3, 101–146, DOI 10.1016/S0022-4049(96)00099-0. MR1600277

[172] H. Gluck, *Almost all simply connected closed surfaces are rigid*, Geometric topology (Proc. Conf., Park City, Utah, 1974), Lecture Notes in Math, vol. 438, Springer, Berlin, 1975, pp. 225–239. MR0400239

[173] G. Gnacadja, *Reachability, persistence, and constructive chemical reaction networks (part II): a formalism for species composition in chemical reaction network theory and application to persistence*, J. Math. Chem. **49** (2011), no. 10, 2137–2157, DOI 10.1007/s10910-011-9896-2. MR2846708

[174] G. Gnacadja, *Reachability, persistence, and constructive chemical reaction networks (part III): a mathematical formalism for binary enzymatic networks and application to persistence*, J. Math. Chem. **49** (2011), no. 10, 2158–2176, DOI 10.1007/s10910-011-9895-3. MR2846709

[175] G. H. Golub and C. F. Van Loan, *Matrix computations*, 4th ed., Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, Baltimore, MD, 2013. MR3024913

[176] R. Goldman, *Polar forms in geometric modeling and algebraic geometry*, Topics in algebraic geometry and geometric modeling, Contemp. Math., vol. 334, Amer. Math. Soc., Providence, RI, 2003, pp. 3–24, DOI 10.1090/conm/334/05972. MR2039963

[177] D. Grayson and M. Stillman, *Macaulay2, a software system for research in algebraic geometry*, `http://www.math.uiuc.edu/Macaulay2/`.

[178] E. Gross, H. A. Harrington, Z. Rosen, and B. Sturmfels, *Algebraic systems biology: a case study for the Wnt pathway*, Bull. Math. Biol. **78** (2016), no. 1, 21–51, DOI 10.1007/s11538-015-0125-1. MR3452313

[179] E. Gross, B. Davis, K. Ho, D. Bates, and H. Harrington, *Numerical algebraic geometry for model selection and its application to the life sciences*, J. R. Soc. Interface **13** (2016), 20160256.

[180] E. Gross and S. Sullivant, *The maximum likelihood threshold of a graph*, Bernoulli **24** (2018), no. 1, 386–407, DOI 10.3150/16-BEJ881. MR3706762

[181] A. Grothendieck and J. Dieudonné, *Éléments de géométrie algébrique I-IV*, Publications Mathématiques de l'IHÉS **4, 8, 11, 17, 20, 24, 28, 32** (1960-67). Second edition of EGA I, Springer, New York, 1971.

[182] J. Gunawardena, *Chemical reaction network theory for in-silico biologists*, lecture notes, Harvard University, 2003, `http://vcp.med.harvard.edu/papers/crnt.pdf`.

[183] J. Gunawardena, *Theory and mathematical methods*, Comprehensive Biophysics **9** (2012), 243–265.

[184] R. Haas, *Characterizations of arboricity of graphs*, Ars Combin. **63** (2002), 129–137. MR1898221

[185] K. Haller, A. Lee-St.John, M. Sitharam, I. Streinu, and N. White, *Body-and-cad geometric constraint systems*, Comput. Geom. **45** (2012), no. 8, 385–405, DOI 10.1016/j.comgeo.2010.06.003. MR2922664

[186] H. Harrington, D. Mehta, H. Byrne and J. Hauenstein, *Decomposing the parameter space of biological networks via a numerical discriminant approach*, 2016, `arXiv:1604.02623[q-bio.MN]`.

[187] J. Harris, *Galois groups of enumerative problems*, Duke Math. J. **46** (1979), no. 4, 685–724. MR552521

[188] J. Harris and L. W. Tu, *On symmetric and skew-symmetric determinantal varieties*, Topology **23** (1984), no. 1, 71–84, DOI 10.1016/0040-9383(84)90026-0. MR721453

[189] V. Hárs and J. Tóth, *On the inverse problem of reaction kinetics*, in *Qualitative Theory of Differential Equations* (M. Farkas, ed.), Colloquia Mathematica Societatis János Bolyai **30**, North-Holland Publ. Co., Amsterdam, 1981, 363–379.

[190] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics, vol. 52, Springer-Verlag, New York-Heidelberg, 1977. MR0463157

[191] J. D. Hauenstein, *Numerically computing real points on algebraic sets*, Acta Appl. Math. **125** (2013), 105–119, DOI 10.1007/s10440-012-9782-3. MR3048642

[192] J. Hauenstein and S. Sherman, *Using monodromy to statistically estimate the number of solutions*, in preparation.

[193] J. D. Hauenstein and A. J. Sommese, *Membership tests for images of algebraic sets by linear projections*, Appl. Math. Comput. **219** (2013), no. 12, 6809–6818, DOI 10.1016/j.amc.2012.12.060. MR3027848

[194] J. D. Hauenstein and A. J. Sommese, *What is numerical algebraic geometry? [Foreword]*. part 3, J. Symbolic Comput. **79** (2017), no. part 3, 499–507, DOI 10.1016/j.jsc.2016.07.015. MR3563094

[195] J. D. Hauenstein and A. J. Sommese, *Witness sets of projections*, Appl. Math. Comput. **217** (2010), no. 7, 3349–3354, DOI 10.1016/j.amc.2010.08.067. MR2733776

[196] J. D. Hauenstein and F. Sottile, *Algorithm 921: alphaCertified: certifying solutions to polynomial systems*, ACM Trans. Math. Software **38** (2012), no. 4, Art. 28, 20, DOI 10.1145/2331130.2331136. MR2972672

[197] J. D. Hauenstein, C. Ikenmeyer, and J. M. Landsberg, *Equations for lower bounds on border rank*, Exp. Math. **22** (2013), no. 4, 372–383, DOI 10.1080/10586458.2013.825892. MR3171099

[198] J. Heintz, B. Kuijpers, and A. Rojas Paredes, *On the intrinsic complexity of elimination problems in effective algebraic geometry*, Recent advances in real complexity and computation, Contemp. Math., vol. 604, Amer. Math. Soc., Providence, RI, 2013, pp. 129–150, DOI 10.1090/conm/604/12071. MR3204156

[199] D. Heldt, M. Kreuzer, S. Pokutta, and H. Poulisse, *Approximate computation of zero-dimensional polynomial ideals*, J. Symbolic Comput. **44** (2009), no. 11, 1566–1591, DOI 10.1016/j.jsc.2008.11.010. MR2561289

[200] J. Hell and A. D. Rendall, *Sustained oscillations in the MAP kinase cascade*, Math. Biosci. **282** (2016), 162–173, DOI 10.1016/j.mbs.2016.10.011. MR3573178

[201] U. Helmke and J. B. Moore, *Optimization and dynamical systems*, Communications and Control Engineering Series, Springer-Verlag London, Ltd., London, 1994. With a foreword by R. Brockett. MR1299725

[202] J. Herzog, A. Simis, and W. V. Vasconcelos, *Approximation complexes of blowing-up rings*, J. Algebra **74** (1982), no. 2, 466–493, DOI 10.1016/0021-8693(82)90034-5. MR647249

[203] J. Herzog, A. Simis, and W. V. Vasconcelos, *Approximation complexes of blowing-up rings. II*, J. Algebra **82** (1983), no. 1, 53–83, DOI 10.1016/0021-8693(83)90173-4. MR701036

[204] O. Hesse, *Über die Bildung der Endgleichung, welche durch Elimination einer Variabeln aus zwei algebraischen Gleichungen hervorgeht, und die Bestimmung ihres Grades* (German), J. Reine Angew. Math. **27** (1844), 1–5, DOI 10.1515/crll.1844.27.1. MR1578380

[205] D. Hilbert, *Ueber die Theorie der algebraischen Formen* (German), Math. Ann. **36** (1890), no. 4, 473–534, DOI 10.1007/BF01208503. MR1510634

[206] D. Hilbert, *Ueber die vollen Invariantensysteme* (German), Math. Ann. **42** (1893), no. 3, 313–373, DOI 10.1007/BF01444162. MR1510781

[207] D. Hilbert, *Mathematische Probleme*, Nachrichten von der Königl. Gesellschaft der Wiss. zu Göttingen, mathematische-physkialische Klasse (1900), 253–297. English translation by M. Newson in *Mathematical Developments Arising from Hilbert Problems*, Proc. Symp. in Pure Math. XXVIII, Part 1, AMS, Providence, RI, 1976, 1–35.

[208] M. Hochster, *Rings of invariants of tori, Cohen-Macaulay rings generated by monomials, and polytopes*, Ann. of Math. (2) **96** (1972), 318–337, DOI 10.2307/1970791. MR304376

[209] K. Holstein, D. Flockerzi, and C. Conradi, *Multistationarity in sequential distributed multisite phosphorylation networks*, Bull. Math. Biol. **75** (2013), no. 11, 2028–2058, DOI 10.1007/s11538-013-9878-6. MR3119394

[210] J. Hong, A. Simis, and W. V. Vasconcelos, *On the homology of two-dimensional elimination*, J. Symbolic Comput. **43** (2008), no. 4, 275–292, DOI 10.1016/j.jsc.2007.10.010. MR2402032

[211] F. Horn, *Necessary and sufficient conditions for complex balancing in chemical kinetics*, Arch. Rational Mech. Anal. **49** (1972/73), 172–186, DOI 10.1007/BF00255664. MR413929

[212] F. Horn, *The dynamics of open reaction systems*, Mathematical aspects of chemical and biochemical problems and quantum chemistry (Proc. SIAM-AMS Sympos. Appl. Math., New York, 1974), Amer. Math. Soc., Providence, R.I., 1974, pp. 125–137. SIAM-AMS Proceedings, Vol. VIII. MR0468667

[213] F. Horn and R. Jackson, *General mass action kinetics*, Arch. Rational Mech. Anal. **47** (1972), 81–116, DOI 10.1007/BF00251225. MR400923

[214] S. Hoşten and S. Ruffa, *Introductory notes to algebraic statistics*, Rend. Istit. Mat. Univ. Trieste **37** (2005), no. 1-2, 39–70 (2006). MR2227048

[215] J. Huh, *Varieties with maximum likelihood degree one*, J. Algebr. Stat. **5** (2014), no. 1, 1–17, DOI 10.18409/jas.v5i1.22. MR3279951

[216] C. Huneke and B. Ulrich, *Divisor class groups and deformations*, Amer. J. Math. **107** (1985), no. 6, 1265–1303 (1986), DOI 10.2307/2374407. MR815763

[217] W. Isaacson, *Leonardo da Vinci*, Simon & Schuster, New York, 2017.

[218] C. G. J. Jacobi, *Theoremata nova algebraica circa systema duarum aequationum, inter duas variabiles propositarum* (Latin), J. Reine Angew. Math. **14** (1835), 281–288, DOI 10.1515/crll.1835.14.281. MR1578080

[219] B. Jackson and J. C. Owen, *Point-line frameworks*, Handbook of geometric constraint systems principles, Discrete Math. Appl. (Boca Raton), CRC Press, Boca Raton, FL, 2019, pp. 487–504. MR3837067

[220] A. Jensen and J. Yu, *Computing tropical resultants*, J. Algebra **387** (2013), 287–319, DOI 10.1016/j.jalgebra.2013.03.031. MR3056698

[221] G. Jeronimo, T. Krick, J. Sabia, and M. Sombra, *The computational complexity of the Chow form*, Found. Comput. Math. **4** (2004), no. 1, 41–117, DOI 10.1007/s10208-002-0078-2. MR2035410

[222] G. Jeronimo and J. Sabia, *Sparse resultants and straight-line programs*, J. Symbolic Comput. **87** (2018), 14–27, DOI 10.1016/j.jsc.2017.05.005. MR3744337

[223] B. Joshi and A. Shiu, *Which small reaction networks are multistationary?*, SIAM J. Appl. Dyn. Syst. **16** (2017), no. 2, 802–833, DOI 10.1137/16M1069705. MR3633777

[224] J. P. Jouanolou, *Singularités rationnelles du résultant* (French), Algebraic geometry (Proc. Summer Meeting, Univ. Copenhagen, Copenhagen, 1978), Lecture Notes in Math., vol. 732, Springer, Berlin, 1979, pp. 183–213. MR555699

[225] J. P. Jouanolou, *Idéaux résultants* (French), Adv. in Math. **37** (1980), no. 3, 212–238, DOI 10.1016/0001-8708(80)90034-1. MR591727

[226] J.-P. Jouanolou, *Le formalisme du résultant* (French), Adv. Math. **90** (1991), no. 2, 117–263, DOI 10.1016/0001-8708(91)90031-2. MR1142904

[227] J.-P. Jouanolou, *Aspects invariants de l'élimination* (French), Adv. Math. **114** (1995), no. 1, 1–174, DOI 10.1006/aima.1995.1042. MR1344713

[228] J.-P. Jouanolou, *Résultant anisotrope, compléments et applications* (French), Electron. J. Combin. **3** (1996), no. 2, Research Paper 2, approx. 91. The Foata Festschrift. MR1392487

[229] J. P. Jouanolou, *Formes d'inertie et résultant: un formulaire* (French), Adv. Math. **126** (1997), no. 2, 119–250, DOI 10.1006/aima.1996.1609. MR1442307

[230] T. Józefiak, A. Lascoux, and P. Pragacz, *Classes of determinantal varieties associated with symmetric and antisymmetric matrices*, Proceedings of the Week of Algebraic Geometry (Bucharest, 1980), Teubner-Texte zur Math., vol. 40, Teubner, Leipzig, 1981, pp. 106–108. MR712519

[231] M. M. Kapranov, B. Sturmfels, and A. V. Zelevinsky, *Chow polytopes and general resultants*, Duke Math. J. **67** (1992), no. 1, 189–218, DOI 10.1215/S0012-7094-92-06707-X. MR1174606

[232] A. Kehrein, M. Kreuzer, and L. Robbiano, *An algebraist's view on border bases*, Solving polynomial equations, Algorithms Comput. Math., vol. 14, Springer, Berlin, 2005, pp. 169–202, DOI 10.1007/3-540-27357-3_4. MR2161988

[233] A. G. Khovanskiĭ, *Fewnomials*, Translations of Mathematical Monographs, vol. 88, American Mathematical Society, Providence, RI, 1991. Translated from the Russian by Smilka Zdravkovska. MR1108621

[234] A. G. Khovanskii and L. Monin, *The resultant of developed systems of Laurent polynomials*, Mosc. Math. J. **17** (2017), no. 4, 717–740, DOI 10.17323/1609-4514-2016-16-4-717-740. MR3734660

[235] C. Király and S.-i. Tanigawa, *Rigidity of body-bar-hinge frameworks*, Handbook of geometric constraint systems principles, Discrete Math. Appl. (Boca Raton), CRC Press, Boca Raton, FL, 2019, pp. 435–459. MR3837065

[236] S. L. Kleiman, *Bertini and his two fundamental theorems*, Rend. Circ. Mat. Palermo (2) Suppl. **55** (1998), 9–37. Studies in the history of modern mathematics, III. MR1661859

[237] D. Kosta and K. Kubjas, *Maximum likelihood estimation of symmetric group-based models via numerical algebraic geometry*, Bull. Math. Biol. **81** (2019), no. 2, 337–360, DOI 10.1007/s11538-018-0523-2. MR3902904

[238] R. Krasauskas, *Toric surface patches*, Adv. Comput. Math. **17** (2002), no. 1-2, 89–113, DOI 10.1023/A:1015289823859. Advances in geometrical algorithms and representations. MR1902537

[239] R. Krasauskas and R. Goldman, *Toric Bézier patches with depth*, Topics in algebraic geometry and geometric modeling, Contemp. Math., vol. 334, Amer. Math. Soc., Providence, RI, 2003, pp. 65–91, DOI 10.1090/conm/334/05976. MR2039967

[240] M. Kreuzer and H. Poulisse, *Subideal border bases*, Math. Comp. **80** (2011), no. 274, 1135–1154, DOI 10.1090/S0025-5718-2010-02432-9. MR2772116

[241] M. Kreuzer and L. Robbiano, *Computational linear and commutative algebra*, Springer, Cham, 2016. MR3559741

[242] M. Kreuzer, H. Poulisse, and L. Robbiano, *From oil fields to Hilbert schemes*, Approximate commutative algebra, Texts Monogr. Symbol. Comput., SpringerWienNewYork, Vienna, 2009, pp. 1–54, DOI 10.1007/978-3-211-99314-9_1. MR2641155

[243] T. Krick, *Straight-line programs in polynomial equation solving*, Foundations of computational mathematics: Minneapolis, 2002, London Math. Soc. Lecture Note Ser., vol. 312, Cambridge Univ. Press, Cambridge, 2004, pp. 96–136. MR2189629

[244] L. Kronecker, *Grundzüge einer arithmetischen Theorie der algebraische Grössen* (German), J. Reine Angew. Math. **92** (1882), 1–122, DOI 10.1515/crll.1882.92.1. MR1579896

[245] E. Kunz, *Residues and duality for projective algebraic varieties*, University Lecture Series, vol. 47, American Mathematical Society, Providence, RI, 2008. With the assistance of and contributions by David A. Cox and Alicia Dickenstein. MR2464546

[246] A. R. Kustin, *The minimal free resolutions of the Huneke-Ulrich deviation two Gorenstein ideals*, J. Algebra **100** (1986), no. 1, 265–304, DOI 10.1016/0021-8693(86)90078-5. MR839583

[247] A. Kustin, C. Polini, and B. Ulrich, *The bi-graded structure of symmetric algebras with applications to Rees rings*, J. Algebra **469** (2017), 188–250, DOI 10.1016/j.jalgebra.2016.08.014. MR3563012

[248] A. R. Kustin, C. Polini, and B. Ulrich, *The equations defining blowup algebras of height three Gorenstein ideals*, Algebra Number Theory **11** (2017), no. 7, 1489–1525, DOI 10.2140/ant.2017.11.1489. MR3697146

[249] P. Lairez, *A deterministic algorithm to compute approximate roots of polynomial systems in polynomial average time*, Found. Comput. Math. **17** (2017), no. 5, 1265–1292, DOI 10.1007/s10208-016-9319-7. MR3709332

[250] G. Laman, *On graphs and rigidity of plane skeletal structures*, J. Engrg. Math. **4** (1970), 331–340, DOI 10.1007/BF01534980. MR0269535

[251] J. M. Landsberg, *Geometry and complexity theory*, Cambridge Studies in Advanced Mathematics, vol. 169, Cambridge University Press, Cambridge, 2017. MR3729273

[252] D. Lazard, *Algèbre linéaire sur $K[X_1, \cdots, X_n]$, et élimination* (French), Bull. Soc. Math. France **105** (1977), no. 2, 165–190. MR0491702

[253] D. Lazard, *Résolution des systèmes d'équations algébriques* (French, with English summary), Theoret. Comput. Sci. **15** (1981), no. 1, 77–110, DOI 10.1016/0304-3975(81)90064-5. MR619687

[254] A. Lee-St.John and J. Sidman, *Combinatorics and the rigidity of CAD systems*, Comput.-Aided Des. **45** (2013), no. 2, 473–482, DOI 10.1016/j.cad.2012.10.030. MR3041222

[255] G. Lecerf, *Kronecker: a Magma package for polynomial system solving*, `lecerf.perso.math.cnrs.fr/software/kronecker/distribution.html`.

[256] A. K. Lenstra, H. W. Lenstra Jr., and L. Lovász, *Factoring polynomials with rational coefficients*, Math. Ann. **261** (1982), no. 4, 515–534, DOI 10.1007/BF01457454. MR682664

[257] J. Limbeck, *Computation of Approximate Border Bases and Applications*, PhD thesis, University of Passau, 2014.

[258] A. Leykin, *Numerical algebraic geometry*, J. Softw. Algebra Geom. **3** (2011), 5–10, DOI 10.2140/jsag.2011.3.5. MR2881262

[259] K.-N. Lin, *Cohen-Macaulayness of Rees algebras of diagonal ideals*, J. Commut. Algebra **6** (2014), no. 4, 561–586, DOI 10.1216/JCA-2014-6-4-561. MR3294862

[260] K.-N. Lin and C. Polini, *Rees algebras of truncations of complete intersections*, J. Algebra **410** (2014), 36–52, DOI 10.1016/j.jalgebra.2014.03.022. MR3201047

[261] D. Lindsey, *Digital Gehry*, Birkhäuser, Basel, 2001.

[262] L. Lovász and Y. Yemini, *On generic rigidity in the plane*, SIAM J. Algebraic Discrete Methods **3** (1982), no. 1, 91–98, DOI 10.1137/0603009. MR644960

[263] T. Luu Ba, L. Busé and B. Mourrain, *Curve/surface intersection problems by means of matrix representations*, in *Proceedings of the 2009 conference on Symbolic Numeric Computation* (H. Kai and H. Sekigawa, eds.), ACM, New York, 2009, 71–78.

[264] F. S. MacAulay, *Some Formulae in Elimination*, Proc. Lond. Math. Soc. **35** (1903), 3–27, DOI 10.1112/plms/s1-35.1.3. MR1577000

[265] F. Macaulay, *The Algebraic Theory of Modular Systems*, Cambridge Univ. Press, Cambridge, 1916.

[266] D. Maclagan and B. Sturmfels, *Introduction to tropical geometry*, Graduate Studies in Mathematics, vol. 161, American Mathematical Society, Providence, RI, 2015. MR3287221

[267] J. Madsen, *Equations of Rees algebras of ideals in two variables*, `arXiv:1511.04073[math.AC]`.

[268] J. C. Maxwell. *On the calculation of the equilibrium and stiffness of frames* Philos. Mag. **27** (1864), 294–299.
[269] H. Melenk, H. Möller and W. Neun, *Symbolic solution of large stationary chemical kinetics problems*, IMPACT Comput. Sci. Eng. **1** (1989), 138–167.
[270] K. Menger, *New Foundation of Euclidean Geometry*, Amer. J. Math. **53** (1931), no. 4, 721–745, DOI 10.2307/2371222. MR1506850
[271] F. Mertens, *Über die bestimmenden Eigenschaften der Resultante von n Formen mit n Veränderlichen*, Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der Kaiserlichen Akademie der Wissenschaften, II. Abtheilung **93** (1886), 527–566.
[272] F. Mertens, *Zur Theorie der Elimination*, Teil II, Sitzungsber. Akad. Wien **108** (1899), 1344–1386.
[273] F. Meyer, *Zur Theorie der reducibeln ganzen Functionen von n Variabeln* (German), Math. Ann. **30** (1887), no. 1, 30–74, DOI 10.1007/BF01564528. MR1510432
[274] E. Miller and B. Sturmfels, *Combinatorial commutative algebra*, Graduate Texts in Mathematics, vol. 227, Springer-Verlag, New York, 2005. MR2110098
[275] Ferd. Minding, *Ueber die Bestimmung des Grades einer durch Elimination hervorgehenden Gleichung* (German), J. Reine Angew. Math. **22** (1841), 178–183, DOI 10.1515/crll.1841.22.178. English translation with commentary by D. Cox and M. Rojas in *Topics in Algebraic Geometry and Geometric Modeling* (R. Goldman and R. Krasauskas, eds.) MR1578276
[276] S. Morey and B. Ulrich, *Rees algebras of ideals with low codimension*, Proc. Amer. Math. Soc. **124** (1996), no. 12, 3653–3661, DOI 10.1090/S0002-9939-96-03470-3. MR1343713
[277] S. Müller, E. Feliu, G. Regensburger, C. Conradi, A. Shiu, and A. Dickenstein, *Sign conditions for injectivity of generalized polynomial maps with applications to chemical reaction networks and real algebraic geometry*, Found. Comput. Math. **16** (2016), no. 1, 69–97, DOI 10.1007/s10208-014-9239-3. MR3451424
[278] S. Müller and G. Regensburger, *Generalized mass action systems: complex balancing equilibria and sign vectors of the stoichiometric and kinetic-order subspaces*, SIAM J. Appl. Math. **72** (2012), no. 6, 1926–1947, DOI 10.1137/110847056. MR3024168
[279] S. Müller, J. Hofbauer, and G. Regensburger, *On the bijectivity of families of exponential/generalized polynomial maps*, SIAM J. Appl. Algebra Geom. **3** (2019), no. 3, 412–438, DOI 10.1137/18M1178153. MR3992044
[280] E. Netto, *Vorlesungen über Algebra*, Volume II, Teubner, Leipzig, 1900.
[281] E. Netto and R. Le Vavasseur, *Fonctions rationnelles*, in *Encyclopédie des sciences mathématiques pures et appliquées* Tome I, Volume 2 (J. Molk, ed.), Gauthiers-Villars, Paris and B. G. Teubner, Leipzig, 1909, 1–232.
[282] I. Newton, *The mathematical papers of Isaac Newton. Vol. II: 1667–1670*, Edited by D. T. Whiteside, with the assistance in publication of M. A. Hoskin, Cambridge University Press, London-New York, 1968. MR0228320
[283] D. G. Northcott, *Finite free resolutions*, Cambridge Tracts in Mathematics, vol. 71, Cambridge University Press, Cambridge-New York-Melbourne, 1976. MR0460383
[284] N. Obatake, A. Shiu, X. Tang and A. Torres, *Oscillations and bistability in a model of ERK regulation*, `arXiv:1903.02617[q-bio.MN]`.
[285] A. Patel and S. Shvartsman, *Outstanding questions in developmental ERK signaling*, Development **145** (2018), dev143818.
[286] E. Penchèvre, *L'élimination en algèbre aux XVII*$^{\text{e}}$ *et XVIII*$^{\text{e}}$ *siècles* (French, with French summary), Historia Sci. (2) **14** (2004), no. 2, 101–117. MR2145854
[287] E. Penchèvre, *Histoire de la théorie de l'élimination*, Ph.D. Thesis, l'Université Paris VII, 2006.
[288] E. Penchèvre, *Etienne Bézout on elimination theory*, 2016, `arXiv:1606.03711[math.HO]`.
[289] E. Penchèvre, *La théorie arithmétique des grandeurs algébriques de Kronecker* (*1882*), 2018, `arXiv:1801.04327[math.HO]`.
[290] M. Pérez Millán and A. Dickenstein, *The structure of MESSI biological systems*, SIAM J. Appl. Dyn. Syst. **17** (2018), no. 2, 1650–1682, DOI 10.1137/17M1113722. MR3811780
[291] M. Pérez Millán, A. Dickenstein, A. Shiu, and C. Conradi, *Chemical reaction systems with toric steady states*, Bull. Math. Biol. **74** (2012), no. 5, 1027–1065, DOI 10.1007/s11538-011-9685-x. MR2909119

[292] M. Plecnik and R. Fearing, *Finding only finite roots to large kinematic synthesis systems*, J. Mech. Robotics **9** (2017), 021005.

[293] S.-D. Poisson, *Mémoire sur l'élimination dans les équations algébriques*, J. Éc. polytech. Math. **4** (1802), 199–203. Reproduced in [**287**, Appendix B].

[294] H. Pollaczek-Geiringer, *Über die Gliederung ebener Fachwerke*, Z. Angew. Math. Mech. **7** (1927), 58–72.

[295] M. Rojas, Review of [**164**], Mathematical Reviews, MR2178092 (2006m:13026).

[296] T.-Y. Li, J. M. Rojas, and X. Wang, *Counting real connected components of trinomial curve intersections and m-nomial hypersurfaces*, Discrete Comput. Geom. **30** (2003), no. 3, 379–414, DOI 10.1007/s00454-003-2834-8. MR2002964

[297] A. Rosen, J. Sidman and L. Theran, *Algebraic matroids in action*, `arXiv:1809.00865[math.CO]`.

[298] B. Roth, *Rigid and flexible frameworks*, Amer. Math. Monthly **88** (1981), no. 1, 6–21, DOI 10.2307/2320705. MR619413

[299] F. Rouillier, M.-F. Roy, and M. Safey El Din, *Finding at least one point in each connected component of a real algebraic set defined by a single equation*, J. Complexity **16** (2000), no. 4, 716–750, DOI 10.1006/jcom.2000.0563. MR1801591

[300] B. Rubinstein, H. Mattingly, A. Berezhkovskii and S. Shvartsman, *Long-term dynamics of multisite phosphorylation*, Mol. Biol. Cell. **27** (2017), 2331–2340.

[301] A. Sadeghimanesh and E. Feliu, *Gröbner bases of reaction networks with intermediate species*, Adv. in Appl. Math. **107** (2019), 74–101, DOI 10.1016/j.aam.2019.02.006. MR3921073

[302] A. Sadeghimanesh and E. Feliu, *The multistationarity structure of networks with intermediates and a binomial core network*, Bull. Math. Biol. **81** (2019), no. 7, 2428–2462, DOI 10.1007/s11538-019-00612-1. MR3977620

[303] M. Sáez, C. Wiuf and E. Feliu, *Nonnegative linear elimination for chemical reaction networks*, 2018, `arXiv:1807.00061[q-bio.MN]`.

[304] M. Safey El Din, *RAGlib library*, `https://www-polsys.lip6.fr/~safey/RAGLib/`.

[305] G. Salmon, *Traité de Géométrie analytique a trois dimensiones*, Paris, Gauthier-Villars, 1882.

[306] N. Schappacher, *A historical sketch of B. L. van der Waerden's work in algebraic geometry: 1926–1946*, Episodes in the history of modern algebra (1800–1950), Hist. Math., vol. 32, Amer. Math. Soc., Providence, RI, 2007, pp. 245–283. MR2353499

[307] M. Schatzman, *Numerical Analysis*, Cambridge Univ. Press, Cambridge, 2002.

[308] H. Schenck, A. Seceleanu, and J. Validashti, *Syzygies and singularities of tensor product surfaces of bidegree* $(2,1)$, Math. Comp. **83** (2014), no. 287, 1337–1372, DOI 10.1090/S0025-5718-2013-02764-0. MR3167461

[309] I. J. Schoenberg, *Remarks to Maurice Fréchet's article "Sur la définition axiomatique d'une classe d'espace distanciés vectoriellement applicable sur l'espace de Hilbert" [MR1503246]*, Ann. of Math. (2) **36** (1935), no. 3, 724–732, DOI 10.2307/1968654. MR1503248

[310] H. Schubert, *Kalkül der abzählenden Geometrie* (German), Springer-Verlag, Berlin-New York, 1979. Reprint of the 1879 original; With an introduction by Steven L. Kleiman. MR555576

[311] T. Sederberg, *Computer Aided Geometric Design*, 2016, `tom.cs.byu.edu/~557/text/cagd.pdf`.

[312] T. Sederberg, D. Anderson and R. Goldman, *Implicit representation of parametric curves and surfaces*, Comput. Vision Graphics Image Process. **28** (1984), 72–84.

[313] T. Sederberg, R. Goldman, and H. Du, *Implicitizing rational curves by the method of moving algebraic curves*, J. Symbolic Comput. **23** (1997), no. 2-3, 153–175, DOI 10.1006/jsco.1996.0081. Parametric algebraic curves and applications (Albuquerque, NM, 1995). MR1448692

[314] T. Sederberg and F. Chen, *Implicitization using moving curves and surfaces*, in *SIGGRAPH 95 Proceedings* (R. Cook, ed.), ACM, New York, 1995, 301–308.

[315] T. W. Sederberg, T. Saito, D. X. Qi, and K. S. Klimaszewski, *Curve implicitization using moving lines*, Comput. Aided Geom. Design **11** (1994), no. 6, 687–706, DOI 10.1016/0167-8396(94)90059-0. MR1305914

[316] A. Seidenberg, *A new decision method for elementary algebra*, Ann. of Math. (2) **60** (1954), 365–374, DOI 10.2307/1969640. MR63994

[317] B. Servatius, H. Servatius, and J. Graver, *Combinatorial rigidity*, Graduate Studies in Mathematics, vol. 2, American Mathematical Society, Providence, RI, 1993. MR1251062

[318] F. Severi, *Sul principio della conservazione del numero*, Rend. Circ. Mat. Palermo **33** (1912), 313–327.

[319] L.-Y. Shen and R. Goldman, *Strong μ-bases for rational tensor product surfaces and extraneous factors associated to bad base points and anomalies at infinity*, SIAM J. Appl. Algebra Geom. **1** (2017), no. 1, 328–351, DOI 10.1137/16M1091952. MR3672369

[320] G. Shinar and M. Feinberg, *Structural sources of robustness in biochemical reaction networks*, Science **327(5971)** (2010), 1389–1391.

[321] A. Shiu and T. de Wolff, *Nondegenerate multistationarity in small reaction networks*, Discrete Contin. Dyn. Syst. Ser. B **24** (2019), no. 6, 2683–2700. MR3960599

[322] J. Sidman and A. St. John, *The rigidity of frameworks: theory and applications*, Notices Amer. Math. Soc. **64** (2017), no. 9, 973–978, DOI 10.1090/noti1575. MR3699772

[323] S. Smale, *Mathematical problems for the next century*, Math. Intelligencer **20** (1998), no. 2, 7–15, DOI 10.1007/BF03025291. MR1631413

[324] A. J. Sommese and C. W. Wampler II, *The numerical solution of systems of polynomials*: *Arising in engineering and science*, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2005. MR2160078

[325] A. J. Sommese, J. Verschelde, and C. W. Wampler, *Numerical irreducible decomposition using projections from points on the components*, Symbolic computation: solving equations in algebra, geometry, and engineering (South Hadley, MA, 2000), Contemp. Math., vol. 286, Amer. Math. Soc., Providence, RI, 2001, pp. 37–51, DOI 10.1090/conm/286/04753. MR1874270

[326] A. J. Sommese, J. Verschelde, and C. W. Wampler, *Symmetric functions applied to decomposing solution sets of polynomial systems*, SIAM J. Numer. Anal. **40** (2002), no. 6, 2026–2046 (2003), DOI 10.1137/S0036142901397101. MR1974173

[327] F. Sottile, *Toric ideals, real toric varieties, and the moment map*, Topics in algebraic geometry and geometric modeling, Contemp. Math., vol. 334, Amer. Math. Soc., Providence, RI, 2003, pp. 225–240, DOI 10.1090/conm/334/05984. MR2039975

[328] G. Staglianò, *A package for computations with classical resultants*, J. Softw. Algebra Geom. **8** (2018), 21–30, DOI 10.2140/jsag.2018.8.21. MR3820371

[329] H. J. Stetter, *Numerical polynomial algebra*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2004. MR2048781

[330] A. St. John, *Generic rigidity of body-and-cad frameworks*, Handbook of geometric constraint systems principles, Discrete Math. Appl. (Boca Raton), CRC Press, Boca Raton, FL, 2019, pp. 505–524. MR3837068

[331] V. Strassen, *Gaussian elimination is not optimal*, Numer. Math. **13** (1969), 354–356, DOI 10.1007/BF02165411. MR0248973

[332] B. Sturmfels, *Algorithms in invariant theory*, 2nd ed., Texts and Monographs in Symbolic Computation, SpringerWienNewYork, Vienna, 2008. MR2667486

[333] B. Sturmfels, *Gröbner bases and convex polytopes*, University Lecture Series, vol. 8, American Mathematical Society, Providence, RI, 1996. MR1363949

[334] B. Sturmfels, *Solving systems of polynomial equations*, CBMS Regional Conference Series in Mathematics, vol. 97, Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2002. MR1925796

[335] B. Sturmfels, *Introduction to resultants*, Applications of computational algebraic geometry (San Diego, CA, 1997), Proc. Sympos. Appl. Math., vol. 53, Amer. Math. Soc., Providence, RI, 1998, pp. 25–39, DOI 10.1090/psapm/053/1602347. MR1602347

[336] B. Sturmfels, *Can biology lead to new theorems?*, Clay Mathematics Institute Annual Report 2005, Feature Article,13, 22–26.

[337] B. Sturmfels and S. Sullivant, *Toric ideals of phylogenetic invariants*, J. Computational Biology **12** (2005), 204–228.

[338] B. Sturmfels and J. Tevelev, *Elimination theory for tropical varieties*, Math. Res. Lett. **15** (2008), no. 3, 543–562, DOI 10.4310/MRL.2008.v15.n3.a14. MR2407231

[339] B. Sturmfels, J. Tevelev, and J. Yu, *The Newton polytope of the implicit equation* (English, with English and Russian summaries), Mosc. Math. J. **7** (2007), no. 2, 327–346, 351, DOI 10.17323/1609-4514-2007-7-2-327-346. MR2337885

[340] J. Sylvester, *A Method of determining by mere inspection the derivatives from two equations of any degree*, Philos. Mag. **XVI** (1840), 132-135.

[341] J. Sylvester, *On a theory of the syzygetic relations of two rational integral functions, comprising an application to the theory of Sturm's functions, and that of the greatest algebraical common measure*, Philosophical Transactions **143** (1853), 407–548.

[342] T.-S. Tay, *Rigidity of multigraphs. I. Linking rigid bodies in n-space*, J. Combin. Theory Ser. B **36** (1984), no. 1, 95–112, DOI 10.1016/0095-8956(84)90016-9. MR742389

[343] S. Telen and M. Van Barel, *A stabilized normal form algorithm for generic systems of polynomial equations*, J. Comput. Appl. Math. **342** (2018), 119–132, DOI 10.1016/j.cam.2018.04.021. MR3808461

[344] S. Telen, B. Mourrain and M. Van Barel, *Solving polynomial systems via a stabilized representation of quotient algebras*, 2017, `arXiv:1711.04543[math.AG]`.

[345] M. Thomson and J. Gunawardena, *The rational parameterisation theorem for multisite post-translational modification systems*, J. Theoret. Biol. **261** (2009), no. 4, 626–636, DOI 10.1016/j.jtbi.2009.09.003. MR2974153

[346] M. Thomson and J. Gunawardena, *Unlimited multistability in multisite phosphorylation systems*, Nature **460(7252)** (2009), 274–277.

[347] W. Trinks, *Über B. Buchbergers Verfahren, Systeme algebraischer Gleichungen zu lösen* (German, with English summary), J. Number Theory **10** (1978), no. 4, 475–488, DOI 10.1016/0022-314X(78)90019-7. MR515056

[348] E. Tschirnhaus, *Methodus auferendi omnes terminos intermedios ex data equatione*, Acta Eruditorium **2** (1683), 204–207. English Translation *A method for removing all intermediate terms from a given equation* by R. Green, ACM SIGSAM Bulletin **37**, March 2003, 1–3.

[349] C. Uhler, *Geometry of maximum likelihood estimation in Gaussian graphical models*, Ann. Statist. **40** (2012), no. 1, 238–261, DOI 10.1214/11-AOS957. MR3014306

[350] B. L. van der Waerden, *Zur Nullstellentheorie der Polynomideale* (German), Math. Ann. **96** (1927), no. 1, 183–208, DOI 10.1007/BF01209162. MR1512314

[351] B. van der Waerden, *Ein algebraisches Kriterium für die Lösbarkeit von homogenen Gleichungen*, Nederl. Akad. Wetensch. Proc. **29** (1926), 142–149.

[352] B. L. van der Waerden, *Der Multiplizitätsbegriff der algebraischen Geometrie* (German), Math. Ann. **97** (1927), no. 1, 756–774, DOI 10.1007/BF01447893. MR1512387

[353] B. L. van der Waerden, *The foundation of algebraic geometry from Severi to André Weil*, Arch. History Exact Sci. **7** (1971), no. 3, 171–180, DOI 10.1007/BF00357215. MR1554142

[354] C. van Loan, *Introduction to Scientific Computing: A Matrix-Vector Approach Using MATLAB*, Second Edition, Pearson, Upper Saddle River, NJ, 2000.

[355] C. W. Wampler and A. J. Sommese, *Numerical algebraic geometry and algebraic kinematics*, Acta Numer. **20** (2011), 469–567, DOI 10.1017/S0962492911000067. MR2805156

[356] C. Wampler, A. Morgan and A. Sommese, *Complete solution of the nine-point path synthesis problem for four-bar linkages*, J. Mech. Des. **114** (1992), 153–159.

[357] L. Wang and E. D. Sontag, *On the number of steady states in a multiple futile cycle*, J. Math. Biol. **57** (2008), no. 1, 29–52, DOI 10.1007/s00285-007-0145-z. MR2393207

[358] E. Waring, *Meditationes algebraicæ*, American Mathematical Society, Providence, RI, 1991. Translated from the Latin, edited and with a foreword by Dennis Weeks; With an appendix by Franz X. Mayer, translated from the German by Weeks. MR1146921

[359] A. Weil, *Foundations of Algebraic Geometry*, American Mathematical Society Colloquium Publications, vol. 29, American Mathematical Society, New York, 1946. MR0023093

[360] J. Weyman, *Cohomology of vector bundles and syzygies*, Cambridge Tracts in Mathematics, vol. 149, Cambridge University Press, Cambridge, 2003. MR1988690

[361] N. White and W. Whiteley, *The algebraic geometry of motions of bar-and-body frameworks*, SIAM J. Algebraic Discrete Methods **8** (1987), no. 1, 1–32, DOI 10.1137/0608001. MR872054

[362] N. L. White and W. Whiteley, *The algebraic geometry of stresses in frameworks*, SIAM J. Algebraic Discrete Methods **4** (1983), no. 4, 481–511, DOI 10.1137/0604049. MR721619

[363] C. Wiuf and E. Feliu, *Power-law kinetics and determinant criteria for the preclusion of multistationarity in networks of interacting species*, SIAM J. Appl. Dyn. Syst. **12** (2013), no. 4, 1685–1721, DOI 10.1137/120873388. MR3116636

[364] W. T. Wu, *On the decision problem and the mechanization of theorem-proving in elementary geometry*, Sci. Sinica **21** (1978), no. 2, 159–172. MR0479996

[365] J. Verschelde, *Algorithm 795: PHCpack: A general-purpose solver for polynomial systems by homotopy continuation*, ACM Trans. Math. Software **25** (1999), 251–276.

[366] Wolfram Research, Inc., *Mathematica*, Version 11.1, Champaign, IL, 2017.

[367] O. Zariski, *Generalized weight properties of the resultant of $n + 1$ polynomials in n indeterminates*, Trans. Amer. Math. Soc. **41** (1937), no. 2, 249–265, DOI 10.2307/1989621. MR1501900

[368] O. Zariski and P. Samuel, *Commutative Algebra*, Volume II, Van Nostrand, Princeton, NJ, 1960. Reprint by Springer, New York, 1976.

[369] J. Zheng and T. W. Sederberg, *A direct approach to computing the μ-basis of planar rational curves*, J. Symbolic Comput. **31** (2001), no. 5, 619–629, DOI 10.1006/jsco.2001.0437. MR1828707

[370] G. M. Ziegler, *Lectures on polytopes*, Graduate Texts in Mathematics, vol. 152, Springer-Verlag, New York, 1995. MR1311028

[371] S. Zubė, *Bidegree* $(2, 1)$ *parametrizable surfaces in projective* 3*-space* (English, with English and Lithuanian summaries), Liet. Mat. Rink. **38** (1998), no. 3, 379–402, DOI 10.1007/BF02465903; English transl., Lithuanian Math. J. **38** (1998), no. 3, 291–308 (1999). MR1657885

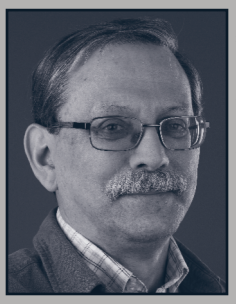
Index

SELECTED PUBLISHED TITLES IN THIS SERIES

134 **David A. Cox,** Applications of Polynomial Systems, 2020
133 **Daniel S. Freed,** Lectures on Field Theory and Topology, 2019
132 **J.M. Landsberg,** Tensors: Asymptotic Geometry and Developments 2016–2018, 2019
131 **Nalini Joshi,** Discrete Painlevé Equations, 2019
130 **Alice Guionnet,** Asymptotics of Random Matrices and Related Models, 2019
129 **Wen-Ching Winnie Li,** Zeta and L-functions in Number Theory and Combinatorics, 2019
128 **Palle E.T. Jorgensen,** Harmonic Analysis, 2018
127 **Avner Friedman,** Mathematical Biology, 2018
126 **Semyon Alesker,** Introduction to the Theory of Valuations, 2018
125 **Steve Zelditch,** Eigenfunctions of the Laplacian on a Riemannian Manifold, 2017
124 **Huaxin Lin,** From the Basic Homotopy Lemma to the Classification of C^*-algebras, 2017
123 **Ron Graham and Steve Butler,** Rudiments of Ramsey Theory, Second Edition, 2015
122 **Carlos E. Kenig,** Lectures on the Energy Critical Nonlinear Wave Equation, 2015
121 **Alexei Poltoratski,** Toeplitz Approach to Problems of the Uncertainty Principle, 2015
120 **Hillel Furstenberg,** Ergodic Theory and Fractal Geometry, 2014
119 **Davar Khoshnevisan,** Analysis of Stochastic Partial Differential Equations, 2014
118 **Mark Green, Phillip Griffiths, and Matt Kerr,** Hodge Theory, Complex Geometry, and Representation Theory, 2013
117 **Daniel T. Wise,** From Riches to Raags: 3-Manifolds, Right-Angled Artin Groups, and Cubical Geometry, 2012
116 **Martin Markl,** Deformation Theory of Algebras and Their Diagrams, 2012
115 **Richard A. Brualdi,** The Mutually Beneficial Relationship of Graphs and Matrices, 2011
114 **Mark Gross,** Tropical Geometry and Mirror Symmetry, 2011
113 **Scott A. Wolpert,** Families of Riemann Surfaces and Weil-Petersson Geometry, 2010
112 **Zhenghan Wang,** Topological Quantum Computation, 2010
111 **Jonathan Rosenberg,** Topology, C^*-Algebras, and String Duality, 2009
110 **David Nualart,** Malliavin Calculus and Its Applications, 2009
109 **Robert J. Zimmer and Dave Witte Morris,** Ergodic Theory, Groups, and Geometry, 2008
108 **Alexander Koldobsky and Vladyslav Yaskin,** The Interface between Convex Geometry and Harmonic Analysis, 2008
107 **Fan Chung and Linyuan Lu,** Complex Graphs and Networks, 2006
106 **Terence Tao,** Nonlinear Dispersive Equations, 2006
105 **Christoph Thiele,** Wave Packet Analysis, 2006
104 **Donald G. Saari,** Collisions, Rings, and Other Newtonian N-Body Problems, 2005
103 **Iain Raeburn,** Graph Algebras, 2005
102 **Ken Ono,** The Web of Modularity: Arithmetic of the Coefficients of Modular Forms and q-series, 2004
101 **Henri Darmon,** Rational Points on Modular Elliptic Curves, 2004
100 **Alexander Volberg,** Calderón-Zygmund Capacities and Operators on Nonhomogeneous Spaces, 2003
99 **Alain Lascoux,** Symmetric Functions and Combinatorial Operators on Polynomials, 2003
98 **Alexander Varchenko,** Special Functions, KZ Type Equations, and Representation Theory, 2003
97 **Bernd Sturmfels,** Solving Systems of Polynomial Equations, 2002

For a complete list of titles in this series, visit the
AMS Bookstore at **www.ams.org/bookstore/cbmsseries/**.

Systems of polynomial equations can be used to model an astonishing variety of phenomena. This book explores the geometry and algebra of such systems and includes numerous applications. The book begins with elimination theory from Newton to the twenty-first century and then discusses the interaction between algebraic geometry and numerical computations, a subject now called numerical algebraic geometry. The final three chapters discuss applications to geometric modeling, rigidity theory, and chemical reaction networks in detail. Each chapter ends with a section written by a leading expert.



Credit: Maria Stenzel @ Amherst College

Examples in the book include oil wells, HIV infection, phylogenetic models, four-bar mechanisms, border rank, font design, Stewart-Gough platforms, rigidity of edge graphs, Gaussian graphical models, geometric constraint systems, and enzymatic cascades. The reader will encounter geometric objects such as Bézier patches, Cayley-Menger varieties, and toric varieties; and algebraic objects such as resultants, Rees algebras, approximation complexes, matroids, and toric ideals. Two important subthemes that appear in multiple chapters are toric varieties and algebraic statistics. The book also discusses the history of elimination theory, including its near elimination in the middle of the twentieth century.

The main goal is to inspire the reader to learn about the topics covered in the book. With this in mind, the book has an extensive bibliography containing over 350 books and papers.

ISBN 978-1-4704-5137-0

9 781470 451370

CBMS/134

For additional information and updates on this book, visit
www.ams.org/bookpages/cbms-134

AMS
www.ams.org